

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

TALEP TAHMİNİ İÇİN MODEL TOPLULUKLARININ KULLANILMASI

YÜKSEK LİSANS TEZİ

İrem İŞLEK

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

ARALIK 2015

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

TALEP TAHMİNİ İÇİN MODEL TOPLULUKLARININ KULLANILMASI

YÜKSEK LİSANS TEZİ

**İrem İŞLEK
(504131524)**

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

Tez Danışmanı: Doç. Dr. Şule ÖĞÜDÜCÜ

ARALIK 2015

İTÜ, Fen Bilimleri Enstitüsü'nün 504131524 numaralı Yüksek Lisans Öğrencisi İrem İŞLEK, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “Talep Tahmini İçin Model Topluluklarının Kullanılması” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Doç. Dr. Şule ÖĞÜDÜCÜ**
İstanbul Teknik Üniversitesi

Jüri Üyeleri : **Doç. Dr. Pınar YOLUM**
Boğaziçi Üniversitesi

Doç. Dr. Gözde ÜNAL
İstanbul Teknik Üniversitesi

Teslim Tarihi : **27 Kasım 2015**
Savunma Tarihi : **30 Aralık 2015**

Anneme ve babama,

ÖNSÖZ

Bu tez kapsamında yapılan çalışma, Bilim, Sanayi ve Teknoloji Bakanlığı SANTEZ programı 0484.STZ.2013-2 kodlu proje kapsamında desteklenmiştir.

Her kararında yanımda olan aileme ve arkadaşlarıma, lisans ve yüksek lisans boyunca bana yol gösteren danışman hocam Doç. Dr. Şule Gündüz Öğüdücü'ye, üzerimde emeği bulunan tüm İTÜ Bilgisayar ve Bilişim Fakültesi hocalarıma teşekkürü borç bilirim.

Kasım 2015

İrem İşlek
(Bilg. Müh.)

İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	vii
İÇİNDEKİLER	ix
ÇİZELGE LİSTESİ.....	xi
ŞEKİL LİSTESİ.....	xiii
ÖZET.....	xv
SUMMARY... ..	xvii
1. GİRİŞ	1
2. LİTERATÜR ÇALIŞMASI	5
3. KULLANILAN YÖNTEMLER	9
3.1 İki Parçalı Çizge Demetleme.....	9
3.2 Yığılmış Genelleme.....	11
3.3 Bayes Ağları	12
3.5 Çok Katmanlı Algılayıcı	12
3.5 Lineer Regresyon	13
3.6 Ardışık Minimal Optimizasyon.....	13
4. ÖNERİLEN YÖNTEM	15
4.1 Veri Kümesini Oluşturma	15
4.2 Veri Hazırlama	15
4.3 İki Parçalı Çizge Oluşturma ve Demetleme	18
4.4 Makine Öğrenmesi Modellerini Oluşturma	19
5. DENEYLER VE SONUÇLAR	24
5.1 Veri Kümesi	25
5.2 Deneysel Sonuçlar.....	25
6. SONUÇ VE ÖNERİLER.....	31
KAYNAKLAR	33
ÖZGEÇMİŞ.....	37

ÇİZELGE LİSTESİ

Sayfa

Çizelge 5.1 :Gruplar için ayrı modeller oluşturma ile tek model oluşturma ile kıyaslaması.	26
Çizelge 5.2 : 4 farklı modelin karşılaştırılması.....	27
Çizelge 5.3 : Yığılmış genelleme.....	28
Çizelge 5.4 : İkili kombinasyonlarla yığılmış genelleme karşılaştırması.....	28
Çizelge 5.5 : Üçlü kombinasyonlarla yığılmış genelleme karşılaştırması.....	28
Çizelge 5.6 : Önerilen yöntemle karşılaştırma	29

ŞEKİL LİSTESİ

	<u>Sayfa</u>
Şekil 1.1 : Doğrudan tedarik zinciri.....	1
Şekil 1.2 : Genişletilmiş tedarik zinciri.	1
Şekil 1.3 : Üst düzey tedarik zinciri.....	2
Şekil 1.4 : Çalışmadaki tedarik zinciri.....	3
Şekil 3.1 : İki parçalı çizge demetleme.....	10
Şekil 3.2 : Yığılmış genelleme yöntemi.	11
Şekil 4.1 : İki parçalı çizgenin çalışmada örnek kullanımı.....	19
Şekil 4.2 : Uygulanan bağımsız yöntemler.....	20
Şekil 4.3 : Çalışmada uygulanan yığılmış genelleme.	21
Şekil 4.4 : İkili kombinasyonlarla yığılmış genelleme.	22
Şekil 4.5 : Üçlü kombinasyonlarla yığılmış genelleme.	23
Şekil 4.6 : Çalışmada önerilen yöntem.	23

TALEP TAHMİNİ İÇİN MODEL TOPLULUKLARININ KULLANILMASI

ÖZET

Tedarik zinciri, ürün, bilgi, servis ve paranın kaynaktan müşteriye veya müşteriden kaynağa aktarımı ve bu aktarım sırasında gerçekleşen tüm süreçleri ifade etmektedir. Depolar için talep tahmini ise herhangi bir deponun ileri tarihli bir zaman dilimi için hangi üründen hangi miktarda satacağını öngörebilmek için bir model geliştirilmesi ve bu model kullanılarak ilgili sonuçların üretilmesi sürecidir. Talep tahmini, tedarik zinciri yönetimi içerisindeki bir alt süreç olup oldukça karmaşık, zor ancak bir o kadar da önemli bir problem olarak kabul edilmektedir. Talep tahmininin karmaşık bir problem olmasındaki en önemli sebeplerden biri oldukça fazla sayıda dış faktörden etkilenmesidir. Bir depo için talep tahmini yapılırken deponun bulunduğu bölge, bu bölgedeki müşterilerin ekonomik durumları, bölgedeki depo sayısı, bölgenin büyüklüğü gibi birçok faktörün göz önünde bulundurulması gerekmektedir. Tüm bunlara ek olarak depo ve ürün sayısının fazlalığı da problemin karmaşıklığını arttırmaktadır. Bu çalışmada yukarıda önemi ve zorluğu vurgulanan, depolar için talep tahmini üzerine çalışılmıştır.

Yapılan çalışmada kullanılan veri kümesi Türkiye’de hizmet veren ulusal bir kuruyemiş firmasının üç yıllık gerçek satış verileri kullanılarak oluşturulmuştur. Oluşturulan veri kümesinde 98 ana dağıtım deposu ve 70 farklı ürün yer almaktadır.

Çalışmada öncelikle seçilen dört farklı makine öğrenmesi algoritmasıyla (Lineer Regresyon, Çok Katmanlı Algılayıcı, Bayes Ağı, Ardışık Minimal Optimizasyon) bağımsız modeller oluşturulmuştur. Ardından çeşitli çizge ve veri madenciliği yaklaşımları kullanılarak depoların ürün satış davranışlarına ilişkin gruplar oluşturulmuştur. Oluşturulan bu gruplar için ayrı ayrı modeller oluşturulduğunda başarımın arttığı görülmüştür.

Bunun ardından Lineer Regresyon, Çok Katmanlı Algılayıcı, Bayes Ağı, Ardışık Minimal Optimizasyon algoritmalarından oluşturulan modellerin model topluluğu oluşturma (model birleştirme) yaklaşımından yararlanılarak birlikte kullanılması denenmiştir. Böyle bir yaklaşımın denenmesinin nedeni, talep tahmini gibi karmaşık bir problemi çözmek için tek bir modelin yetersiz kalmasıdır. Model topluluğu oluşturma yöntemleriyle başarımın artırılabilceği düşünülmüştür. Burada model topluluğu oluşturulması için öncelikle yığılmış genelleme yönteminin kullanıldığı denemeler yapılmıştır. Bu denemeler sırasında ilk olarak yukarıda bahsedilen 4 farklı algoritma birlikte kullanılarak, ardından bu algoritmaların ikili ve üçlü kombinasyonları kullanılarak yığılmış genelleme modelleri oluşturulmuş ve karşılaştırılmıştır. Son olarak modelleri birleştirmek için bu çalışmada önerilen yöntemle başka bir deneme yapılmıştır. Çalışmada önerilen model, yığılmış genelleme yönteminden yola çıkılarak oluşturulmuştur. Ancak söz konusu modelin genelleme modeli için giriş öznitelikleri seçilmesi kısmı yığılmış genelleme yönteminden ayrılmaktadır. Önerilen yöntem kullanılarak tahmin modelinin başarımının arttığı gözlenmiştir.

USING ENSEMBLES OF CLASSIFIERS FOR DEMAND FORECASTING

SUMMARY

Supply chain means all the processes during transferring finance, product, service and information from source to customer or customer to source. There are several types of supply chain models which are composed according to their complexities. Supply chain types can be collected into three main groups: Direct Supply Chain, Extended Supply Chain and Ultimate Supply Chain. Demand forecasting which is a research topic in machine learning is the process of estimating the product quantity that customers will purchase. Demand forecasting for warehouses means constructing a model to estimate demands of products for a future time interval.

In Direct Supply Chain model, the products of a manufacturer are sent to distribution warehouses. Customers reach these products through distribution warehouses of the company. In Extended Supply Chain model, numerous different components are added to supply chain model likewise suppliers of supplier, customers of customers etc. Due to the fact that there are several components are added in this model, it is more complicated than Direct Supply Chain model. In Ultimate Supply Chain model, there are additional components such as third party logistic company, financial providers, market research companies etc. In this model, these components should also be considered during processes.

Demand forecasting which is a sub process of supply chain management is a quite complex and important problem. Because of the complexity of demand forecasting is that it is effected from numerous different factors. For instance, region size, sale power of region community, product type, population of the region should be considered in order to forecast demand of a warehouse. In addition to that, when the number of warehouses and products increase, the demand forecasting problem becomes more and more complex.

Since both the number of warehouses and the variety of products increase in today's competitive and dynamic business environment, accurate demand forecasting becomes more important. Thanks to accurate demand forecasting, planning the other phases of supply chain management can be done in the correct way. In addition to that, the improvement of the accuracy of demand forecasting provides quite significant savings. For instance, accurate demand forecasting provides avoiding various expenditure likewise unnecessary logistic costs, storage costs, redundant producer goods etc. Thus, it is aimed to increase the profitability of company.

The main purpose of this study is developing a model which provides estimating the quantity of sale amounts of different products for main distribution warehouses correctly. In this study, we focus on a Direct Supply Chain model which contains numerous main distribution warehouses. Moreover, every main distribution warehouse contains several sub distribution warehouses in this model. Customers can reach products through sub distribution warehouses. In our problem, customers are end sale points such as supermarkets, canteens, grocery stores etc.

In this thesis, we studied forecasting the demand of main distribution warehouses with low error rate problem for a company. The dataset of this study was constructed from the real sales transaction data of a national dried fruits and nuts company from Turkey. This company gives service to nearly all cities of Turkey. Sales transaction of 2011, 2012 and 2013 were used for this dataset. This company has ninety eight main distribution warehouses. The company produces and distributes seventy different products.

It becomes more difficult to estimate the demand accurately using traditional methods with the increasing number of the variety of products and size of warehouses. Because of this situation, most of the previous studies are focus on limited count of warehouses or products. In addition to that, most of the previously proposed methods are interested in forecasting demand of customers. Differently from this, we focus on forecasting the demand of main distribution warehouses in our study. Moreover, the estimation accuracy of previous studies are not sufficient. Contributions of this thesis can be summarized as follows: proposed methodology can handle numerous main distribution warehouses and products, bipartite graphs and special data mining techniques for bipartite graphs are used for defining sale behaviours of main distribution warehouses, a new methodology for ensembling classifiers is proposed.

There are four basic steps of the methodology. In the first step, dataset is constructed from sale invoices of 2011, 2012 and 2013 years for the national dried fruit and nut company. Constructed dataset contains ninety eight main distribution warehouses and seventy different products. Moreover, the dataset comprises numerous additional information likewise main distribution warehouse information (locations, number of sub distribution warehouses, number of vehicles, number of customers, number of employees, size of sale region etc.), product information (product ontology is used for that), sale amounts and sales time information.

Second step of the methodology is called data preparation. This step is used for cleaning and preparing data for the data mining methodologies. For this reason, dataset is analized and it is found that there are quite small quantities of columns which are empty. Rows which contain empty columns are deleted. After that, a product ontology which is used for defining products and their properties is constructed. This ontology will be used for forecasting demand of a new product. Because of the fact that a new product doesn't have any past sale data, sale data of the nearest neighbours in the ontology of the new product will be used for demand forecasting. After that, numerous new properties are found for warehouses. For instance, six main categories are constructed for main warehouses, customer lifetime values are calculated for warehouses. Also, special day problem is handled in this step. Because, it is noticed that special days (New Year, Ramadan Holiday, Ramadan Month, Sacrifice Holiday etc.) change product demands more than expected.

Third step of the methodology is used for modeling the sale behaviours of warehouses. Because of this step is that some of the main distribution warehouses have different sale behaviours. Some warehouses give service to wider area and they also have more sub distribution warehouses than others. Moreover, some main warehouses may sell higher profit margin products. For this reason, main distribution warehouses are grouped based on their product sale amounts using bipartite graph. After the bipartite graph is constructed, this graph is clustered using Bipartite Graph Clustering algorithm.

In order to decrease the error rate, another bipartite graph is constructed for modeling sale behaviours of sub distribution warehouses. Purpose of this operation is that, some sub distribution warehouses give service to different areas with different purchase power. Bipartite graph which includes sub distribution warehouses is clustered using Bipartite Graph Clustering algorithm.

In the last step of the methodology, stand-alone models are constructed using Multi Layer Perceptron, Linear Regression, Sequential Minimal Optimization and Bayesian Network algorithms. After that, some clusters are composed according to sales behaviours of main distribution warehouses using graphs and data mining methodologies. Constructing separate models for main distribution clusters provides improvement in results. Then, separate models are trained for every sub distribution warehouse clusters. This trial gives better results than previous trials.

After that, models which are composed using Multi Layer Perceptron, Linear Regression, Sequential Minimal Optimization and Bayesian Network algorithms are combined using classification ensemble approach. Because of using this approach is that stand alone models are unefficient to solve demand forecasting problem. Thus, stacked generalization which is an example of classification ensemble is tried for combining these models in this study.

Stacked generalization methodology contains two basic levels which are level-0 and level-1. In level-0, output results are estimated for every separate models. In level-1, another model which uses output results of level-0 as input attributes is trained. Level-1 is called as generalization phase in this methodology.

In the next trial, stacked generalization models are constructed using binary combinations of given four machine learning algorithms. There are six different combinations of these algorithms and separate stacked generalization models are constructed for every different algorithm combinations. For instance, one of the stacked generalization models contains only Multi Layer Perceptron and Sequential Minimal Optimization in the level-0 of stacked generalization model.

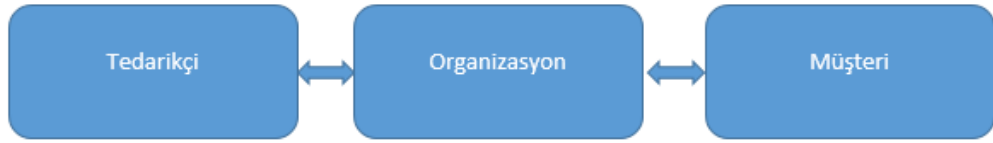
Afterwards, triple combinations of four machine learning algorithms are selected. These four combinations are used for constructing separate stacked generalization models. For example, in one of the trials, level-0 of stacked generalization model contains only Multi Layer Perceptron, Bayesian Network and Linear Regression algorithms. Next trials are done using other triple combinations.

Finally, a new method is applied to combine models. This new method was constructed based on stacked generalization but different attributes were used as inputs of generalization phase (level-1) model. In stacked generalization, level-1 which is called as generalization phase uses results of first step model outputs as input attributes. Differently from stacked generalization, proposed method uses both results of level-0 model outputs and level-0 model input attributes as input attributes of level-1 (generalization model). This methodology decreased the error rate of the system.

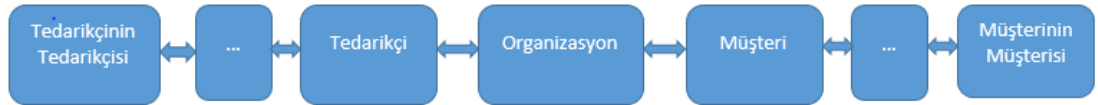
In brief, experiments show that clustering warehouses based on their sale behaviours and training separate models for these clusters provides better results than training one model for all warehouses. Reason of this situation is that, training different models for warehouses which have different sale behaviours is more appropriate for given problem. In addition to that, proposed ensembling classifiers methodology gives better results than stacked generalization methodology.

1. GİRİŞ

Tedarik zinciri, ürün, servis, para ve bilginin kaynaktan müşteriye kadar çift yönlü olarak aktarımı ile ilgili tüm süreçleri kapsayan ve yöneten yapıdır [1]. Tedarik zincirinin ilgili yapının karmaşıklığına göre oluşturulmuş farklı türleri bulunmaktadır. Bu türler temel olarak 3'e ayrılabilir: Doğrudan Tedarik Zinciri, Genişletilmiş Tedarik Zinciri ve Üst Düzey Tedarik Zinciri [2]. Şekil 1.1'de görüldüğü gibi, Doğrudan Tedarik Zinciri'nde üreticinin ürünleri depolara gönderilir ve müşteriler ürünlere bu depolar aracılığıyla ulaşırlar. Genişletilmiş Tedarik Zinciri'nde ise tedarikçinin tedarikçileri, müşterilerin müşterileri vb. bileşenlerin de zinciri dahil edilmesiyle daha karmaşık bir yapı elde edilir. Bu yapı Şekil 1.2'de görülmektedir. Şekil 1.3'teki Üst Düzey Tedarik Zinciri'nde ise Genişletilmiş Tedarik Zinciri'ne üçüncü taraf lojistik firması, finansörler, pazar araştırması yapan firmalar gibi bileşenlerin de eklendiği oldukça karmaşık bir yapı bulunmaktadır.



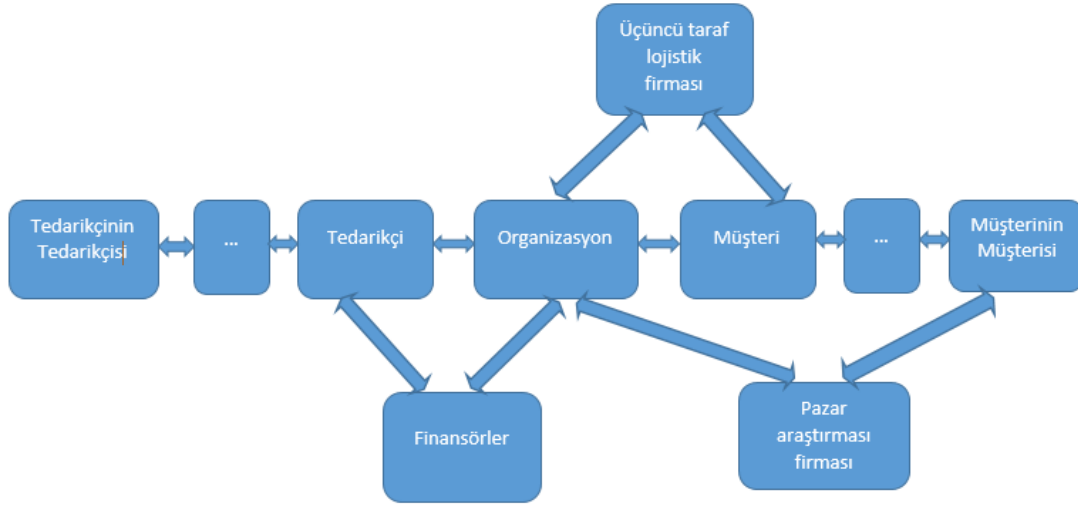
Şekil 1.1: Doğrudan tedarik zinciri.



Şekil 1.2: Genişletilmiş tedarik zinciri.

Talep tahmini ise tedarik zinciri yönetiminin oldukça önemli ve karmaşık bir alt sürecidir. Talep tahmini, tedarik zinciri içerisindeki ilgili bileşenlerin ileri bir zaman diliminde hangi üründen hangi miktarda satın alacaklarını gerçeğe en yakın şekilde tahmin etmek için modeller geliştirilmesi ve bu modeller kullanılarak ilgili miktarların hesaplanması süreci olarak tanımlanabilir. Bu sürecin zorluğunun en temel sebeplerinden biri birçok dış faktörden etkilenmesidir. Örneğin bir ana depo

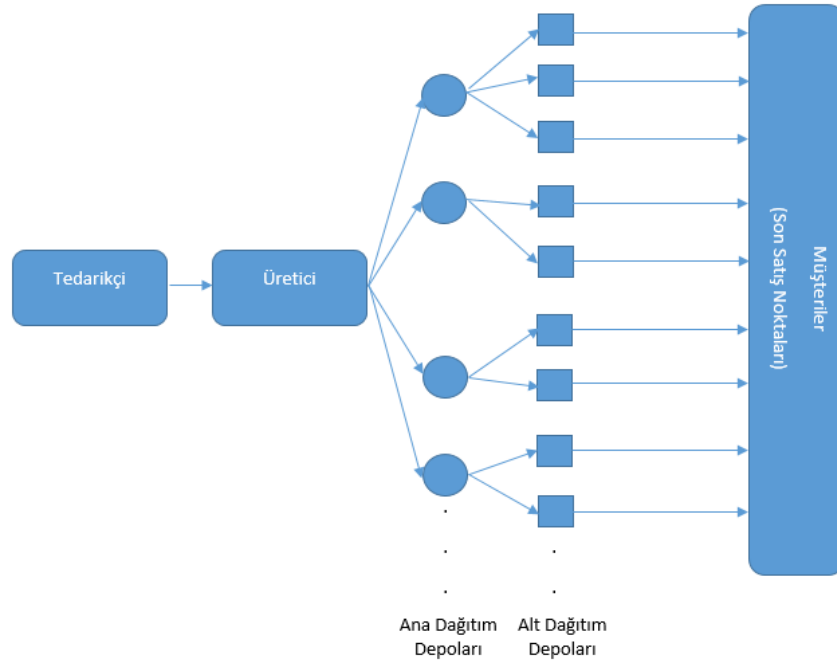
için talep tahmini yapılırken gerçeğe yakın sonuçlar elde edebilmek amacıyla deponun hangi bölgede bulunduğu, bölgenin büyüklüğü, müşteri sayısı, ürün tipi gibi çok çeşitli faktörler göz önünde bulundurulmalıdır.



Şekil 1.3: Üst düzey tedarik zinciri.

Talep tahmininin gerçeğe yakın şekilde yapılması tedarik zincirindeki tüm süreçlerin doğru şekilde planlanabilmesine yardımcı olması açısından oldukça önemlidir. Örneğin talep tahmininin gerçeğe oldukça yakın şekilde yapılmasıyla gereksiz lojistik masrafları, depolama maliyetleri, gereksiz miktarda hammadde tedarik edilmesi gibi oldukça önemli sorunların önüne geçilebilir. Bu sayede firmanın karlılığının artırılması hedeflenmektedir.

İlgili tezin amacı, Doğrudan Tedarik Zinciri yapısına uygun bir firmanın ana dağıtım depolarının hangi zaman diliminde, hangi üründen hangi miktarda satacaklarının gerçeğe en yakın şekilde tahmin edilmesini sağlayabilecek bir model geliştirilmesidir. Tez kapsamında ele alınan tedarik zincirinde bir adet üretici firmanın çok sayıda ana dağıtım deposu, ve her bir ana dağıtım deposunun farklı sayıda alt depoları bulunmaktadır. Bu yapı Şekil 1.4'te görülmektedir. İlgili tedarik zincirinde müşteri olarak kabul edilen kantin, bakkal, market vb. satış noktaları ürünleri alt depolardan temin etmektedirler. Tez kapsamında talep tahmini sonuçları ana dağıtım depoları için hesaplanmaktadır. Tezde kullanılan veri kümesi Türkiye'deki hemen hemen her ile hizmet veren yerel bir kuruyemiş firmasının 2011, 2012 ve 2013 yılları gerçek satış faturaları kullanılarak oluşturulmuştur. Bahsedilen veri kümesinde doksan sekiz farklı ana dağıtım deposu ve yetmiş farklı ürün bulunmaktadır.



Şekil 1.4: Çalışmadaki tedarik zinciri.

İlgili problemde ana dağıtım deposu ve ürün sayısının fazla olması geleneksel yöntemlerin yetersiz kalmasına sebep olmaktadır. Konuyla ilgili yapılmış çalışmaların hemen hemen hepsinde depo sayısı ve ürün sayısı sınırlı tutulmuştur. Ayrıca yapılan çalışmaların çoğunda ana depolar için değil, son satış noktaları (bakkal, market, kantin vb.) için talep tahmini yapılmıştır. Bunun yanı sıra, mevcut yöntemler ile elde edilen başarı oranları yeterince yüksek değildir. Bu tezin konuyla ilgili olarak sağladığı katkılar ise şu şekilde ifade edilebilir:

- Tez kapsamında yapılan çalışmayla fazla sayıda depo ve ürün bulunması durumunda da çözüm sağlayabilecek bir yöntem önerilmiştir.
- Depoların ürün satış davranışlarını belirleyebilmek için çizgeler oluşturulmuş ve bu çizgelerde veri madenciliği yöntemleri uygulanmıştır.
- Oldukça karmaşık olan talep tahmin problemi için tek bir modelin yetersiz kalması sebebiyle birden çok modelin birlikte kullanıldığı bir yöntem önerilmiştir.

Tezin ikinci bölümünde yapılan literatür araştırması, üçüncü bölümünde kullanılan yöntemler, dördüncü bölümde ise önerilen yöntem ve bu yönteme ilişkin detaylar verilmiştir. Tezin beşinci bölümü yapılan deneyler, bu deneylerden elde edilen sonuçların verildiği kısımdır. Altıncı bölümde ise elde edilen sonuçlar açıklanmış, bu sonuçlarla ilgili yorumlardan ve gelecek çalışmalardan bahsedilmiştir.

2. LİTERATÜR ÇALIŞMASI

Ana depolar için talep tahmini, herhangi bir ana deponun ileri tarihli bir zaman diliminde hangi üründen toplamda kaç adet satacağının hesaplanmasıdır. Literatürde talep tahmini için kullanılan yöntemlerin iki ana başlıkta toplandığı görülmüştür. Bunlar tek bir modelden oluşan yöntemler ve birden çok modelin birlikte kullanıldığı hibrid modellerdir. Tek bir modelden oluşan yöntemler de istatistiksel modeller ve makine öğrenmesi modelleri şeklinde ikiye ayrılabilir.

Hareketli ortalama ve Box-Jenkins modelleri istatistiksel modeller içerisinde en çok kullanılan modellere örnek olarak verilebilir. Literatürde talep tahmini için bu yöntemleri kullanan çeşitli çalışmalar mevcuttur. Örneğin bir çalışmada Box-Jenkins sezonsal ARIMA (Autoregressive Integrated Moving Average) modelleri kullanılarak zaman serileri analizi ve tahmini yapılmaya çalışılmıştır [3]. Ancak tedarik zinciri problemi oldukça karmaşık olduğu için bu yöntemler yeterli başarıyı sağlamakta yetersiz kalmışlardır. Buradan yola çıkılarak, geleneksel istatistiksel modeller yerine yapay zeka yöntemlerinin kullanılması fikri ortaya çıkmıştır. Yapılan bir çalışmada istatistiksel yöntemlerin promosyon olmayan dönemler için başarılı sonuç verdiği, ancak promosyon olması durumunda yetersiz kaldığına değinilmiştir. Yine bu çalışmada promosyon durumunda regresyon ağaçlarının geleneksel istatistiksel yöntemlerden daha başarılı olduğu belirtilmektedir [4]. Başka bir çalışmada ise talep tahmini için regresyon ağaçları kullanıldığında, en küçük kare regresyonu, temel bileşen regresyonu, parçalı en küçük kare regresyonu, çarpımsal regresyon ve yarı logaritmik regresyon yöntemlerine kıyasla daha yüksek başarımler elde edildiği vurgulanmıştır [5].

Özellikle yapay sinir ağlarının talep tahminine yönelik birçok çalışmada kullanıldığı görülmüştür [6-11]. Yapay sinir ağlarının geleneksel istatistiksel modellerle kıyaslamasının yapıldığı çeşitli çalışmalar bulunmaktadır. Örneğin Bangladeş'teki bir süpermarketin verileriyle yapılan bir çalışmada öncelikle geleneksel Holt-Winter modeliyle ardından ise yapay sinir ağları ile tahmin yapılmıştır. Hata oranları MAPE (Ortalama Mutlak Yüzdesel Hata) ile ölçülmüş ve yapay sinir ağlarının geleneksel

istatistiksel modellere kıyasla daha yüksek başarımlar sağladığı belirtilmiştir [12]. Yapay sinir ağlarının talep tahmini için yaygın olarak kullanılıyor olması sebebiyle, bu modeli başka modellerle kıyaslayan çeşitli çalışmalar da yapılmıştır. Bunlardan bir tanesinde ANFIS (Adaptive Neural Fuzzy Inference System) ile yapay sinir ağları kıyaslanmıştır. Elde edilen sonuçlarda ANFIS'in kullanılmasıyla daha yüksek başarımlar elde edildiği belirtilmiştir [13]. Başka bir çalışmada ise bir giyim firmasının perakende satış talebi tahmini için tam bağımlı olmayan yapay sinir ağları, tam bağımlı yapay sinir ağları ve SARIMA (Seasonal Autoregressive Integrated Moving Average) yöntemleri karşılaştırılmış ve tam bağımlı olmayan yapay sinir ağlarının diğer iki yöntemle kıyasla daha yüksek başarımlar sağladığı vurgulanmıştır [10]. Talep tahmini için yapay sinir ağlarını lineer regresyon ile kıyaslayan ve test sonuçlarını yedi farklı performans metriği ile karşılaştıran bir çalışmada ise yapay sinir ağlarının daha iyi performans sağladığı belirtilmektedir [14].

Son zamanlarda yapılan çalışmalarda ise başarımları arttırmak amacıyla farklı modellerin birlikte kullanıldığı görülmektedir. Yapılan bir çalışmada öncelikle SARIMA ile birden çok değişkenli lineer regresyon, ardından SARIMA ile dağılım regresyonu birlikte kullanılarak 2 farklı hibrid tahmin yöntemi oluşturulmuştur. Çalışma sonucunda hibrid yöntemlerin tek başına yapay sinir ağları veya SARIMA kullanan yöntemlere kıyasla daha yüksek başarımlar sağladığı belirtilmiştir [15]. Kesikli dalgacık dönüşümü yöntemini yapay sinir ağları ile birlikte kullanan ve bu hibrid yöntem ile ARIMA (Autoregressive Integrated Moving Average)'dan daha düşük hata oranı elde eden bir çalışma da bulunmaktadır [16]. Başka bir çalışmada genetik algoritmalar ile yapay sinir ağları hibrid şekilde kullanılmaktadır. Bu çalışmada günlük süt satış verisi ile çalışılmış olup bahsedilen hibrid yöntem lineer AR (Autoregressive Model) yöntemine kıyasla hatanın yaklaşık %5 azalmasını sağlamıştır [17]. Yine hibrid yaklaşımı kullanan başka bir çalışmada ise Şili'de bir süpermarketin satış verileri ile çalışılmış, ARIMA modeli ve yapay sinir ağlarının birlikte kullanıldığı bir model oluşturulmuştur. Bu modelin hibrid olmayan yöntemlere kıyasla daha düşük hata oranına sahip olduğu vurgulanmıştır [18].

Literatürde birden çok modelden elde edilen tahmin sonuçlarını birleştirerek model toplulukları oluşturma yaklaşımına ilişkin çalışmalara rastlanmaktadır [19,20,21]. Oylama ve ortalama alma gibi basit yöntemlerin bağımsız modellere kıyasla tahmin doğruluğunu arttırdığını gösteren çalışmalar bulunmaktadır [22]. Ancak topluluk

seçme, modelleri birleştirmek için bir fonksiyon eğitme gibi daha karmaşık model topluluğu oluşturma yöntemlerinin kullanılmasıyla daha yüksek doğrulukta sonuç veren model toplulukları elde edildiğini gösteren çalışmalar mevcuttur. Örneğin bu konuda yapılan bir çalışmada yedi test problemi ve on farklı performans metriği kullanılarak yapılan karşılaştırmalar sonucunda model topluluklarının hem bağımsız modellerden hem de ortalama alma gibi basit yaklaşımlı birleştirme yöntemlerinden daha yüksek başarımlar sağladığı belirtilmiştir [23].

Modelleri birleştirmek için kullanılan karmaşık yöntemlerden biri olan yığılmış genellemede ise modelleri birleştirmek için başka bir model kullanılmaktadır [24]. Yığılmış genelleme yöntemi, birden çok öğrenme aşaması içermesi bakımından ardışık öğrenme (cascade learning) yöntemine [25] benzemektedir. Ancak ardışık öğrenme yönteminden farklı olarak yığılmış genellemede sadece iki adet ardışık öğrenme aşaması bulunmaktadır. Bunun yanı sıra, ardışık öğrenmede birden çok temel model olması gibi bir zorunluluk yoktur. Peşpeşe öğrenme modellerinin olması yeterlidir. Yığılmış genellemeyi 21 farklı makine öğrenmesi veri kümesiyle test eden ve bu testler sonucunda başarılı sonuçlar elde edildiğini gösteren bir çalışma da bulunmaktadır [26]. Yığılmış genelleme yöntemini modellerin ayrı ayrı kullanıldığı yöntemlerle kıyaslayan çalışmalar bulunmaktadır. Örneğin bir çalışmada yığılmış genelleme ile bağımsız modeller karşılaştırılmış ve test edilen iki problem için de yığılmış genellemenin daha yüksek başarımlar sağladığı ifade edilmiştir [27]. Buna ek olarak yığılmış genelleme model topluluğu oluşturma yönteminin diğer model topluluğu oluşturma yöntemlerine kıyasla daha yüksek başarımlar sağladığını ifade eden çalışmalar bulunmaktadır. Örneğin, üç farklı öğrenme algoritması kullanılarak oluşturulan yığılmış genelleme modelinin aynı algoritmalar kullanılarak oluşturulan oy çokluğu yöntemine kıyasla üç farklı veri kümesi kullanılarak yapılan testlerde daha iyi performans gösterdiği belirtilmektedir [28].

Yığılmış genelleme çok çeşitli alanlardaki çalışmalarda kullanılmıştır. Protein lokalizasyon tahmini [29], Alzheimer hastalığının erken tanısı [30], hileli finansal durumların tahmini [31], resim sınıflandırma [32], kredi risk değerlendirmesi [33], istenmeyen e-postaların filtrelenmesi [34] gibi çalışmalar bunlara örnektir. Ancak depolar için talep tahmininde yığılmış genellemeyi kullanan bir çalışmaya rastlanmamıştır.

3. KULLANILAN YÖNTEMLER

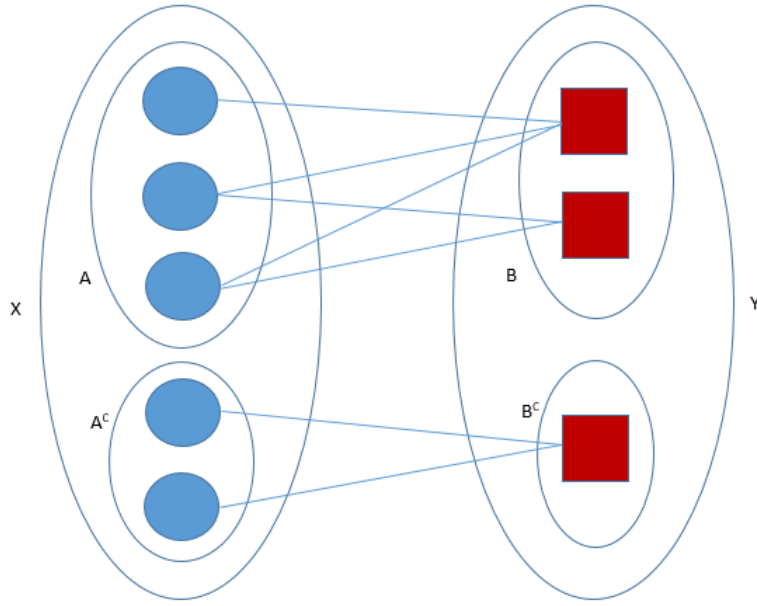
Bu bölümde tez içerisinde kullanılan yöntemlerle ilgili detaylı bilgi verilmektedir. Kullanılan yöntemler İki Parçalı Çizge Demetleme (Bipartite Graph Clustering), Yığılmış Genelleme (Stacked Generalization), Bayes Ağları (Bayesian Networks), Ardışık Minimal Optimizasyon (Sequential Minimal Optimization), Çok Katmanlı Algılayıcı (Multi Layer Perceptron) ve Lineer Regresyon (Linear Regression)'dur.

3.1 İki Parçalı Çizge Demetleme

Bazı uygulamalarda veri $G(X, Y, W)$ şeklinde iki parçalı çizge yapısıyla ifade edilebilmektedir. İki parçalı çizge, (X, Y) şeklinde iki farklı düğüm içeren özelleştirilmiş bir çizgedir. Bu çizge yapısında düğümler arasındaki ağırlıklandırılmış ayrıtlar (W) ile ifade edilir. İki parçalı çizgelerin en karakteristik özelliği aynı tip düğümler arasında ayrıt bulunamamasıdır. Bu çizge türünde yalnızca farklı türdeki düğümler arasında ayrıt bulunabilmektedir.

Bu tezde iki parçalı çizgeler depolar ve sattıkları ürünler arasındaki ilişkiyi modelleyebilmek için kullanılmıştır. Oluşturulan iki parçalı $G(X, Y, W)$ çizgesinde (X, Y) düğüm kümesi sırasıyla depoları ve ürünleri sembolize ederken, (W) ağırlıklandırılmış ayrıtları ise ilgili deponun o üründen o yıl kaç adet sattığı bilgisine karşılık gelmektedir.

Tezde kullanılan iki parçalı çizge demetleme algoritması [35] öncelikle iki parçalı çizgeyi iki farklı iki parçalı çizgeye ayırmaktadır. Ardından elde edilen iki parçalı çizgelerin her biri için ikiye ayırma işlemini rekürsif bir şekilde tekrarlamaktadır. Bu işlem iki parçalı çizgenin ikiye ayıramadığı duruma kadar devam etmektedir. Şekil 3.1'de bir iki parçalı çizgenin $\Pi(A, B)$ bölümlenmesi ile ikiye ayrılması görülmektedir. Bu çizgede, X düğüm kümesi $X = A \cup A^c$ ve Y düğüm kümesi ise $Y = B \cup B^c$ şeklinde alt düğüm kümelerine ayrılmışlardır. Sözü geçen bölümlenmede A düğüm kümesi ve B düğüm kümesi bir demet oluştururken A^c ile B^c düğüm kümeleri ise ikinci bir demeti oluşturmaktadır.



Şekil 3.1: İki parçalı çizge demetleme.

Denklem 3.1’de görüldüğü gibi, çizgenin nereden ikiye ayrılacağıının belirlenmesi için çizge içerisinde eşleşmeyen düğümlerin benzerliklerinin minimum olduğu bir Ncut bölümlemesi aranmaktadır.

$$\min_{\pi(A,B)} Ncut(A, B) \quad (3.1)$$

Bu Ncut bölümlemesinin bulunabilmesi için öncelikle çizge içerisindeki mümkün olan tüm bölümlemelere ilişkin Ncut hesaplamasının denklem 3.2 kullanılarak yapılması gerekmektedir.

$$Ncut(A, B) = \frac{cut(A, B)}{W(A, Y) + W(X, B)} + \frac{cut(A^c, B^c)}{W(A^c, Y) + W(X, B^c)} \quad (3.2)$$

Ncut hesaplamasının yapılabilmesi için denklem 3.3’ten yararlanılması gerekmektedir. Burada $W(A, Y)$ bir ucu A düğümü diğer ucu ise Y düğümü olan tüm ayrıtların ağırlıklarının toplamını ifade etmektedir.

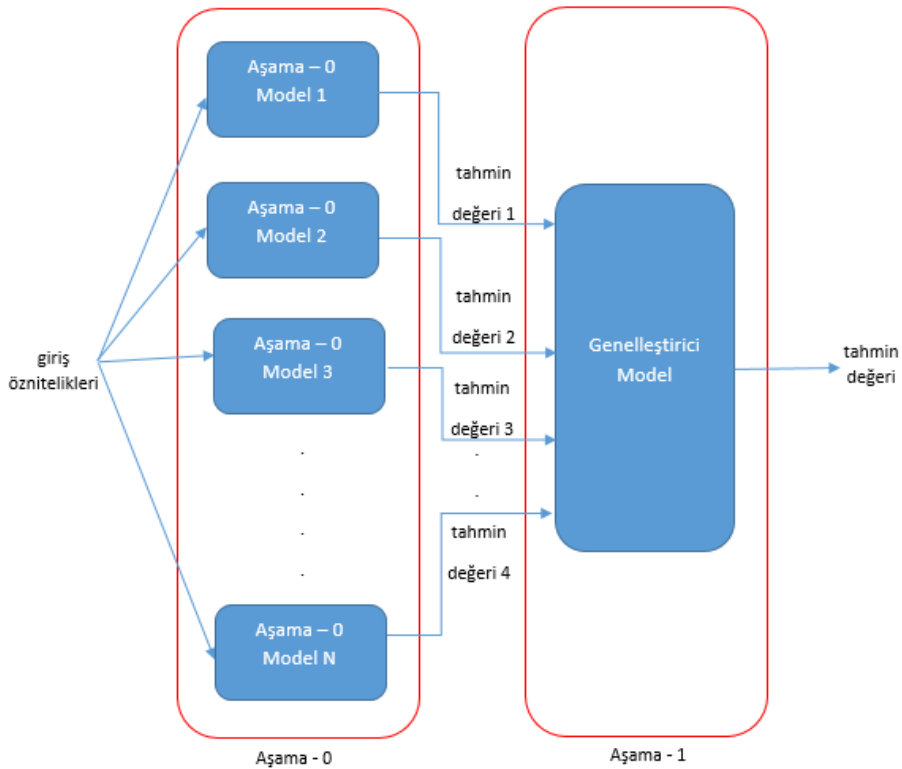
$$cut(A, B) = W(A, B^c) + W(A^c, B) \quad (3.3)$$

İki parçalı çizgeyi demetlemek için bu algoritmanın seçilmesinin sebebi bahsedilen algoritmanın iki parçalı çizgeleri demetlemekte diğer demetleme algoritmalarına (çizgeler için özel olarak geliştirilmemiş olanlara) kıyasla daha başarılı olduğunun belirtilmesidir [35].

3.2 Yığılmış Genelleme

Model toplulukları (ensemble of classifiers) kavramı özetle ayrı ayrı modellerin performansını yükseltmek amacıyla birden çok modelin bir arada kullanılmasıdır. Modelleri bir arada kullanabilmek için iki farklı yöntem izlenebilir: oylama ve yığma. Yığılmış genelleme yaklaşımı, bir veya birden çok modelin hata oranını minimize etmek için kullanılan bir model topluluğu oluşturma (model birleştirme) yöntemidir [24]. Bu yaklaşım David H. Wolpert tarafından 1992 yılında önerilmiştir.

Şekil 3.2’de görülen yığılmış genelleme yaklaşımında ardışık iki temel aşama vardır. İlk aşama olan aşama-0’da birbirinden bağımsız makine öğrenmesi modelleri bulunmaktadır. Bu bağımsız makine öğrenmesi modelleri genelleme aşaması olan aşama-1’de birleştirilmektedir. Aşama-1’de, aşama-0’daki her bir bağımsız makine öğrenmesi modelinden elde edilen sonuçlar ile yeni bir genelleme modeli eğitilmektedir. Genelleme aşamasında kullanılacak olan genelleme modeli için ilgili probleme uygun olduğu düşünülen herhangi bir makine öğrenmesi modeli tercih edilebilmektedir.



Şekil 3.2: Yığılmış genelleme yöntemi.

Yığılmış genellemenin aşama-0 modelleri bir eğitim veri kümesi kullanılarak eğitilir. Ardından aşama-0'daki modellerin her birinden elde edilen sonuçlar kullanılarak yeni bir eğitim kümesi oluşturulur. Bu eğitim veri kümesi ise aşama-1'in eğitiminde kullanılmaktadır. Burada dikkat edilmesi gereken nokta, yeni eğitim kümesi oluşturulurken kullanılan örneklerin aşama-0'daki modellerin eğitimi sırasında kullanılmamış olmasıdır. Aşama-1 genelleme modelini (ve modelin genelini) test edebilmek için üçüncü bir veri kümesine ihtiyaç duyulmaktadır. Öncelikle bu veri kümesindeki her bir örnek için aşama-0'daki bağımsız modellerle sonuçlar elde edilmesi gerekmektedir. Ardından bu model sonuçları aşama-1'deki genelleme modeline girdi olarak verilir ve genelleme modelinden modelin bütünüdür sonucı elde edilir.

3.3 Bayes Ağları

Bayes ağlarının en temel özelliği istatistiksel ağlar olmaları ve düğümler arası geçiş yapan kolların istatistiksel kararlara göre seçiliyor olmasıdır. Bayes ağları yönlü ve dönüşsüz (Directed Acyclic Graph - DAG) çizge yapısındadır. Bu ağlarda düğümler raslantı değişkenlerini, düğümler arası bağlar ise raslantı değişkenleri arasındaki olasılıklı bağımlılık durumlarını ifade etmektedir. Bayes ağlarında her değişken için bir tablo tutulmaktadır. Bayes ağının yapısı ve bu ağdaki tüm değişkenler biliniyorsa Bayes Formülü kullanılarak koşullu olasılıklar denklem 3.4'teki gibi hesaplanmaktadır. Ağ yapısının bilindiği ancak bazı değişkenlerin bilinmediği durumda ise yinelemeli öğrenme yöntemlerinden yararlanılmaktadır [36].

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (3.4)$$

3.4 Çok Katmanlı Algılayıcı

Yapay sinir ağlarının bir örneği olan çok katmanlı algılayıcılar giriş katmanı, gizli katmanlar ve çıkış katmanı olmak üzere üç ana katmandan oluşmaktadır. Bahsedilen yapıda, aktivasyon fonksiyonuna sahip birçok nöron birbirine hiyerarşik olarak bağlıdır. Çok katmanlı algılayıcılarda geri yayılım (back propagation) öğrenme yöntemi kullanılmaktadır.

Aktivasyon fonksiyonu en genel tanımıyla, bir değişkeni başka bir boyuta taşımakta kullanılan doğrusal veya doğrusal olmayan fonksiyon şeklinde ifade edilmektedir. Modelin eğitim hızı aktivasyon fonksiyonu ile ilişkilidir. Yaygın olarak kullanılan üç aktivasyon fonksiyonu denklem 3.5'te görülen sigmoid, denklem 3.6'da görülen hiperbolik tanjant ve denklem 3.7'de görülen adım basamaktır.

$$y = \frac{1}{1 + e^{-net}} \quad (3.5)$$

$$y = \frac{1 - e^{-2net}}{1 + e^{-2net}} \quad (3.6)$$

$$y = \begin{cases} 1 & net \geq 0 \\ -1 & net < 0 \end{cases} \quad (3.7)$$

Geri yayılım öğrenme yönteminde tüm ağırlık değerleri eğim azaltma (gradient descent) yönteminin hata fonksiyonunu minimize etmesiyle bulunmaktadır. Hesaplanan hata değeri, nöronun çıkışından girişine giderken nöronun aktivasyon fonksiyonunun türevine uygulanmak ve hata değeri tüm ağırlıklara giriş değerleri oranında dağıtılmaktadır [37].

3.5 Lineer Regresyon

Lineer regresyon, bağımlı bir değişken ile bir veya birden fazla bağımsız değişken arasındaki ilişkiyi incelemek için kullanılan bir yöntemdir. Regresyon temel olarak tek değişkenli ve birden çok değişkenli olmak üzere ikiye ayrılmaktadır. Tek değişkenli regresyon analizinde bir bağımlı değişken ile bir bağımsız değişken arasındaki ilişki incelenmektedir. Bu doğrultuda, ilgili doğrusal ilişkiyi ifade eden bir doğru denklemi bulunmaya çalışılmaktadır. Çok değişkenli regresyon analizinde ise denklem 3.8'de görüldüğü gibi, birden çok bağımsız değişken ile bağımlı değişken arasındaki ilişki formüle edilmektedir [38].

$$y = \alpha + x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + \dots + x_n\beta_n + \varepsilon \quad (3.8)$$

3.6 Ardışık Minimal Optimizasyon

Ardışık minimal optimizasyon, destek vektör makinelerinin eğitimi sırasında ortaya çıkan ikinci dereceden programlama probleminin çözülmesi için kullanılan bir algoritmadır. Ardışık minimal optimizasyon algoritması, 1998 yılında John Platt tarafından önerilmiştir [39].

Destek vektör makinelerindeki ikinci dereceden programlama problem denklem 3.9'daki gibi ifade edilmektedir. Denklem 3.10'da i 'nin 1'den n 'e kadar tüm değerleri için α_i değerinin 0 ve C değeri aralığında olması gerekmektedir. Ayrıca denklem 3.11'de i 'den n 'e kadar tüm α_i ve y_i değerlerinin çarpımlarının toplamı 0'a eşit olmalıdır.

$$\max_{\alpha} = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j K(x_i x_j) \alpha_i \alpha_j \quad (3.9)$$

$$0 \leq \alpha_i \leq C \quad (3.10)$$

$$\sum_i^n y_i \alpha_i = 0 \quad (3.11)$$

Bu algoritma, destek vektör makinelerinin eğitiminde ortaya çıkan ikinci dereceden programlama problemini iteratif bir şekilde bir dizi daha küçük alt problemlere bölmektedir. Elde edilen küçük alt problemler iki adet Lagrange çarpanı içermektedirler. Problem α_1 'in 0'dan büyük eşit ve α_2 'nin C değerinden küçük eşit olduğu durumda denklem 3.12'deki şekle dönüşmektedir.

$$y_1 \alpha_1 + y_2 \alpha_2 = k \quad (3.12)$$

Bu aşamadan sonra yapılması gereken tek boyutlu ikinci dereceden fonksiyonun minimumunun bulunmasıdır. Bunun için öncelikle optimizasyon problemi için uygun bir α_1 çarpanı seçilmektedir. Ardından bir α_2 çarpanı seçilmekte ve (α_1, α_2) çifti optimize edilmektedir. Bu iki aşama yakınsayana kadar devam etmektedir [39].

4. ÖNERİLEN YÖNTEM

Önerilen yöntemin en temel amacı, gelecek tarihli bir zaman aralığı için hangi deponun hangi üründen kaç adet satacağını öngörebilmektir. Yöntemin ilk aşamasında veri kümesi veri madenciliği çalışmalarını uygulamaya hazır hale getirilmiştir. Ardından iki parçalı çizge demetleme algoritması kullanılarak satış davranışları benzer olan ana ve alt depolar gruplanmıştır. Bu grupların her biri için ayrı ayrı modeller oluşturulmuştur. Bu aşamada modeller oluşturulurken dört farklı makine öğrenmesi algoritması kullanılmıştır. Yöntemin sonraki aşamasında farklı algoritmalar kullanılarak oluşturulan modeller yığılmış genelleme yöntemi kullanılarak birleştirilmiştir. Son aşamada ise yığılmış genelleme yönteminden yola çıkılarak oluşturulan yeni yöntemle modellerin birleştirilmesi sağlanmıştır. Uygulanan tüm adımların detayları bu bölümde açıklanmaktadır.

4.1 Veri Kümesini Oluşturma

Yöntemin ilk aşaması tahmin modeli için gerekli olan verileri içeren ve veri madenciliği uygulamalarına uygun bir veri kümesi oluşturulmasıdır. Veri kümesi, ulusal bir kuruyemiş firmasının 2011, 2012 ve 2013 yılı satış faturalarından oluşturulmuştur. Oluşturulan veri kümesinde 98 adet ana dağıtım deposu ve 70 adet ürün bulunmaktadır. Veri kümesinde deponun lokasyonu, kaç adet alt depoya sahip olduğu, dağıtım araçlarının sayısı, personel sayısı, haftalık sattığı ürün sayıları, kaç metrekaare satış alanına sahip olduğu, müşteri sayısı gibi depoya ilişkin bilgiler ve ürünün kategorisi (çalışmada bunun için ürün ontolojisi oluşturulmuştur), satış miktarları, satışın gerçekleşme zamanı gibi ürüne ilişkin bilgiler yer almaktadır.

4.2 Veri Hazırlama

Veri madenciliği çalışmalarının en önemli adımlarından biri veri hazırlama aşamasıdır. Çalışmanın bu aşamasında veri temizlenmiş ve veri madenciliği yöntemlerini uygulamaya hazır hale getirilmiştir.

Öncelikle veri incelenmiş, boş veya geçersiz kayıtlar belirlenmiştir. Yapılan çalışma sonucunda boş sütun içeren satır sayısının toplam satır sayısına oranla oldukça az miktarda olduğu görülmüştür. Veriyi bozmamak için bu kayıtların silinmesine karar verilmiştir.

Yine bu aşamada, ürünler için bir ürün ontolojisi oluşturulmuştur. Ürün ontolojisi oluşturmaktaki amaç, hem ontoloji vasıtasıyla diğer sistemlere entegre olabilecek bir yapı oluşturabilmek hem de soğuk başlatma (cold start) probleminden kaçınmaktır. Soğuk başlatma, önceki satışlarına ilişkin veri bulunmayan yeni bir ürün için satış talep tahmini yapılması problemi. Bu çalışmada ontoloji oluşturulurken Protege [40] isimli araçtan yararlanılmıştır. Tüm ürünler için belirleyici olan özellikler detaylıca tanımlanmıştır. Elde edilen ontoloji, 4 ana ve 28 alt ürün kategorisi içermektedir. Yetmiş farklı ürün bu ontoloji kullanılarak tanımlanmıştır.

Veride yapılan diğer incelemeler sonucunda herhangi bir ürün için hiç satış yapılmayan hafta sayısının oldukça fazla olduğu görülmüştür. Bu duruma sebep olabilecek 3 farklı durum tespit edilmiştir:

- Ürünün ana depo stoğunda bulunmaması
- Ürünün yeni çıkmış olması ve bu nedenle hiç satılmaması (reklamı yapılmamış olabilir)
- Ürünün talep olmadığı için satılmaması

Belirtilen sebeplerin ilkini sağlayan satırların veri kümesinde kullanılmaması gerektiğine karar verilmiştir. Çünkü belirtilen durumda satış yapılmaması ürünlere talep olmamasından değil, ürünün stoğunun bulunmamasından kaynaklanmaktadır. Bu nedenle bu satırların veri kümesine eklenmesinin talep tahmininde hataya sebep olacağı düşünülmüştür.

Belirtilen sebeplerden ikincisinin gerçekleşmesi durumunda ürünün reklamlarının dönmeye başladığı tarihten itibaren ontolojideki en yakın komşularından yararlanılarak hareketli ortalama değerinin hesaplanmasına karar verilmiştir.

Belirtilen sebeplerden üçüncüsü için ise ek bir çalışma yapılmasına gerek yoktur. Bu satırlar olduğu gibi veri kümesine eklenmelidir.

Ardından, ana dağıtım depoları için mevcut bilgileri kullanarak yeni özellikler elde edilmeye çalışılmıştır. Bunun için öncelikle modelde kullanmak üzere depolar için

sınıflar belirlenmesine karar verilmiştir. Bu doğrultuda öncelikle depoların aylık ortalama satışlarının türk lirası cinsinden değerine göre sınıflandırılmasına karar verilmiştir. Burada elde edilen değerler için bölütleme çalışması yapılmış ve her bir deponun hangi sınıfa dahil olduğu belirlenmiştir. Başka bir özellik oluşturmak için ise öncelikle eldeki veriden depoların mevcut her bir müşterisinin türü çıkarılmıştır. Burada müşterilerin türleri bakkal, süpermarket, benzinlik, kantin, lokanta vb. şeklindedir. Müşteri türlerinin her birine bir ağırlık değeri atanmış ve bu ağırlıklar kullanılarak her bir deponun tüm müşterilerinden bir ortalama müşteri değeri hesaplanmıştır. Buna ek olarak her bir deponun yıllar arasındaki müşteri sayısı, araç sayısı ve personel sayısındaki yüzdesel artışlar da hesaplanmıştır. Bir başka özellik olarak da müşteri yaşam boyu değeri (CLV) hesaplanmıştır. Müşteri yaşam boyu değeri hesaplanırken C_s periyodun başlangıcındaki müşteri sayısı, C_e periyodun bitişindeki müşteri sayısı ve C_n periyot süresince kazanılan yeni müşteri sayısı olacak şekilde denklem 4.1'de görüldüğü şekilde hesaplanmaktadır. Bu denklem kullanılarak elde edilen müşteri yaşam boyu değeri, yıllık kazanç değeri ile çarpılarak başka bir özellik daha oluşturulmuştur.

$$CLV = \frac{1}{((C_e - C_n)/C_s) * 100} \quad (4.1)$$

Veri hazırlama aşamasının bir sonraki adımında ise hareketli ortalama (moving average) hesaplaması yapılmıştır. Hareketli ortalama hesaplaması yapılırken ilgili ürünün o depoda geçmiş haftalarda kaçar adet satıldığı bilgisi kullanılmıştır. Herhangi bir t haftası için hareketli ortalama denklem 4.2'de görüldüğü şekilde hesaplanmaktadır.

$$Hareketli\ ort.(t) = \frac{\sum_{i=1}^{hafta\ sayısı} satış\ miktarı(t-i)}{hafta\ sayısı} \quad (4.2)$$

Bayram, ramazan ayı, yılbaşı vb. özel günlerin talep tahminini önemli ölçüde etkilediği bilinmektedir. Veride yapılan incelemeler sonucunda örneğin yılbaşından önceki iki hafta satışların arttığı, yeni yılın ilk haftasında ise satışların azaldığı görülmüştür. Eldeki veri kümesi kuruyemiş firmasına ait olduğu için maç, konser vb. özel günlerin de satışlardaki etkisinin fazla olduğu farkedilmiştir. Bayram, ramazan ayı, maç, konser vb. özel günlerin her yıl aynı tarihlere denk gelmemesi problemini çözebilmek için esnek bir yapı oluşturulmasına karar verilmiştir. Bu yapıda herhangi bir özel gün eklenmek istendiğinde bu özel güne ilişkin bir tanımlama yapılmalı ve

bu özel güne ilişkin geçmiş yıllardaki örnekler işaretlenmelidir. Eğer tahmin yapılacak hafta özel gün içeriyorsa, o özel gün etiketine sahip geçmiş yıllardaki satış miktarları alınmakta ve model girişinde diğer özelliklerle birlikte bu satış miktarları kullanılmaktadır. Eğer tahmin yapılacak hafta özel bir gün içermiyorsa, geçmiş haftalardaki satış miktarlarıyla hareketli ortalama hesaplanmakta ve model girişinde hareketli ortalama değeri ve diğer özellikler kullanılmaktadır. Bu esnek yapı kullanılarak satışların azalma eğiliminde olduğu bilinen yılın ilk haftası için yılbaşı sonrası gibi bir özel gün tanımı yapılabilir. Buna ek olarak satışların artma eğiliminde olduğu bilinen yılbaşı öncesi ve yılbaşı için de başka bir özel gün tanımlanabilmektedir.

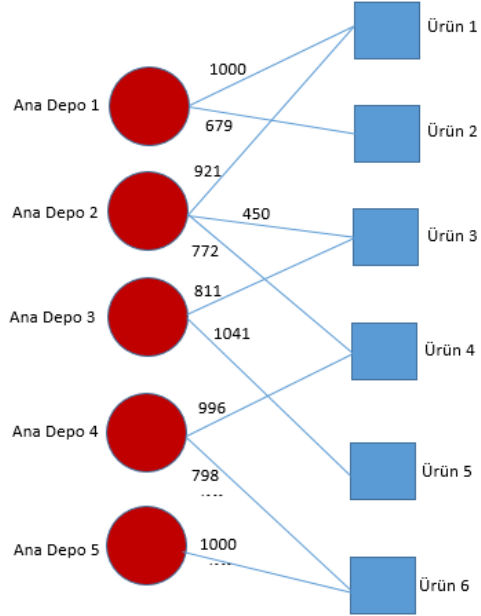
4.3 İki Parçalı Çizge Oluşturma Ve Demetleme

Çalışma sırasında bazı ana dağıtım depolarının diğerlerinden farklı satış davranışı gösterdiği gözlenmiştir. Örneğin, İstanbul'daki ana dağıtım depolarının hizmet alanları diğer depolara kıyasla oldukça geniştir. Buna ek olarak bu depolar oldukça fazla sayıda alt depoya sahiptirler. Depoların hangi bölgeye hizmet verdikleri, bu bölgedeki insanların alım güçleri gibi birçok depo özelliği o deponun sattığı ürünleri etkilemektedir. Örneğin bir deponun müşterileri yüksek fiyatlı ürünleri daha çok tercih ederken başka bir bölgeye hizmet veren bir diğer deponun müşterileri daha ucuz ürünleri tercih edebilir. Bu doğrultuda, ana dağıtım depoları ürün satış miktarlarına göre iki parçalı çizge demetleme kullanılarak gruplanmıştır. Bunun için ana dağıtım depoları ve tüm ürünler kullanılarak iki parçalı çizge oluşturulmuştur. Bu çizgede bir düğüm tipi ana dağıtım depoları, diğer düğüm tipi ise ürünlerdir. Eğer bir depo bir üründen satmışsa bu iki düğüm arasında bir ayrıt bulunmaktadır. Bu ayrıtın ağırlığı ise o deponun o üründen toplamda kaç adet sattığı bilgisidir. Bu iki parçalı çizgenin temsili bir gösterimi Şekil 4.1'de görülmektedir.

Benzer ürün satış eğilimlerine sahip depoları gruplayabilmek için iki parçalı çizge demetleme algoritmasından faydalanılmıştır. Bu algoritma iki parçalı çizgeleri demetlemede geleneksel demetleme algoritmalarına kıyasla daha yüksek başarımla elde ettiği belirtildiği için tercih edilmiştir [35]. Burada 29 farklı ana dağıtım deposu grubu elde edilmiştir.

Bu aşamanın ardından benzer bir çizge oluşturma-demetleme aşaması alt depolar kullanılarak yapılmıştır. Bunun sebebi İstanbul gibi oldukça geniş bölgelere hizmet

vermekte olan ve fazla sayıda alt depo içeren yerlerde alt depolar arası satış davranışlarının da farklılık gösterebilmesidir. Bir önceki aşamada uygulanan yaklaşım ana depolar yerine alt depolar olacak şekilde tekrarlanmış ve böylelikle benzer satış davranışlarına sahip alt depo grupları elde edilmiştir. Burada elde edilen alt depo grubu sayısı 97'dir.



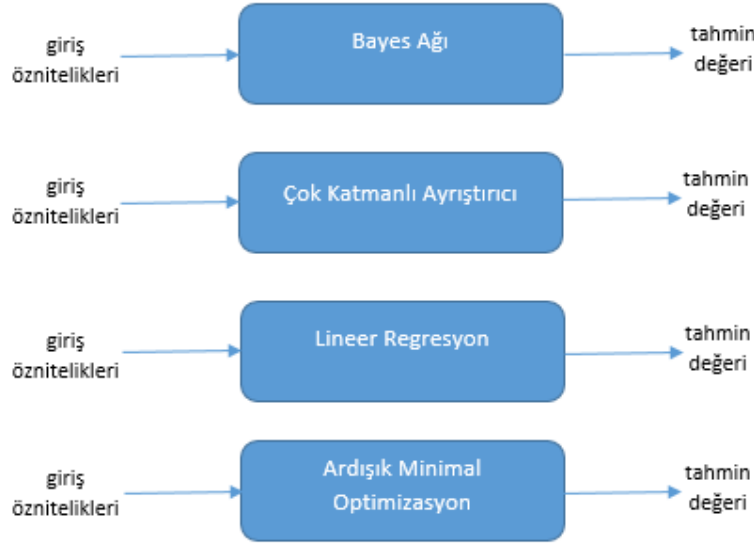
Şekil 4.1: İki parçalı çizgenin çalışmada örnek kullanımı.

4.4 Makine Öğrenmesi Modellerini Oluşturma

Yöntemin bu aşamasında makine öğrenmesi algoritmaları kullanılarak modeller oluşturulmuştur. Modeller oluşturulurken hareketli ortalama değeri, depoya ilgili özellikler, ürünle ilgili özellikler kullanılmıştır. Bayes ağları kullanılarak kurulan bu modelde yukarıda bahsedilen özellikler birer rastlantısal değişkene karşılık gelmektedir. Tahmin sonuçları, bu rastlantısal değişkenlerden yola çıkılarak hesaplanan olasılıklara göre belirlenmektedir. Bayes ağları ile kurulan ilk modelde tüm ana depolar kullanılarak tek bir model oluşturulmuştur.

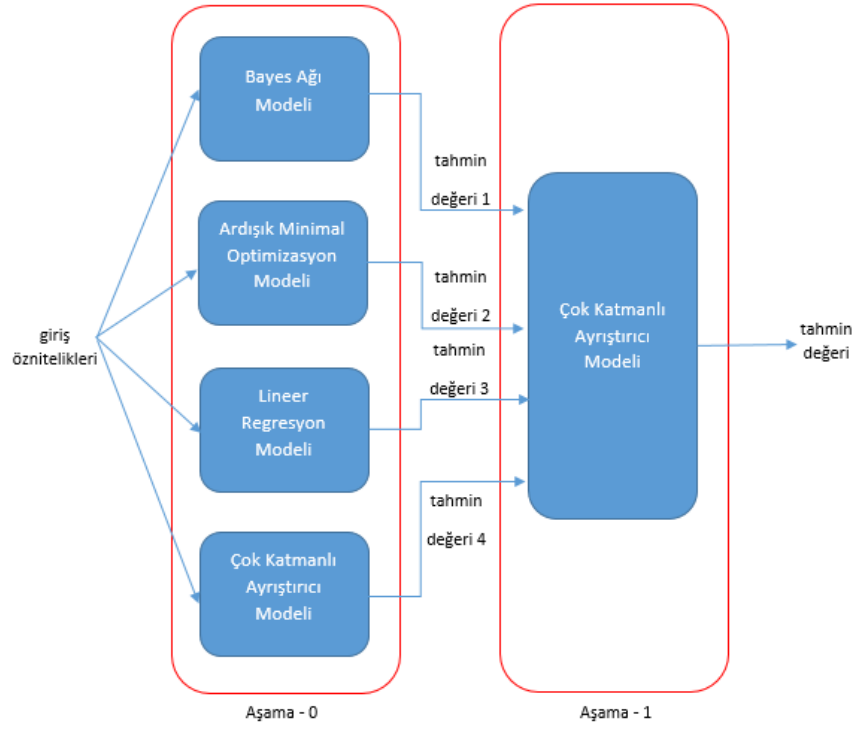
Yapılan ikinci denemede ise iki parçalı çizge demetleme yöntemi ile elde edilen ana dağıtım deposu gruplarının herbiri için ayrı ayrı Bayes ağı modeli oluşturulmuştur. Üçüncü denemede ise alt depo grupları için ayrı ayrı modeller kurulmuştur. Bu üç deneme arasında en yüksek başarımlı alt depo grupları için ayrı ayrı Bayes ağları oluşturulan denemede elde edilmiştir.

Bir sonraki aşamada Bayes ağları seçilen üç makine öğrenmesi algoritmasıyla kıyaslanmıştır. Bu algoritmalar Ardışık Minimal Optimizasyon, Çok Katmanlı Ayrıştırıcılar ve Lineer Regresyon'dur. Bayes ağları için bu aşamaya dek en yüksek başarımlı alt depo grupları için ayrı ayrı modeller oluşturularak elde edildiği için seçilen üç algoritma için de aynı yaklaşım uygulanmıştır. Bu üç algoritmanın her biri tüm alt depo grupları için ayrı ayrı modeller oluşturularak test edilmiştir. Bu yapıya ilişkin bir gösterim Şekil 4.2'de görülmektedir.



Şekil 4.2: Uygulanan bağımsız modeller.

Kullanılan Yöntemler kısmında detayları verilmiş olan yığılmış genelleme, iki farklı aşama içeren ve hata oranını azaltma amacıyla seçilen modelleri birleştiren bir model topluluğu oluşturma (model birleştirme) yöntemidir. Bu çalışmada da yığılmış genelleme aynı amaçla seçilmiş olup, belirlenen dört farklı algoritmayla oluşturulan modelleri birleştirmek için kullanılmaktadır. Çalışmada uygulanan ve Şekil 4.3'te görülen yığılmış genellemenin aşama-0'ında dört farklı algorithmadan oluşturulan modeller yer almaktadır. Aşama-0'daki modellerin her birinden elde edilen tahmin sonuçları sonraki aşamadaki genelleme modeli ile nihai tahmin sonucunun oluşturulmasında kullanılmaktadır. Her bir örnek için aşama-0'daki bağımsız modellerin her birinden bir tahmin sonucu üretilmektedir. Ardından bu tahmin sonuçları aşama-1'deki genelleme modelinin girişine verilmektedir. Bu çalışmada genelleme modeli için Çok Katmanlı Ayrıştırıcı kullanılmıştır. Aşama-1'deki genelleme modeli, aşama-0 modellerinden elde edilen tahmin sonuçlarını gerçek satış miktarlarıyla eşleştirir. Aşama-1'den elde edilen sonuç, tüm yöntemin nihai tahmin sonucunu vermektedir.

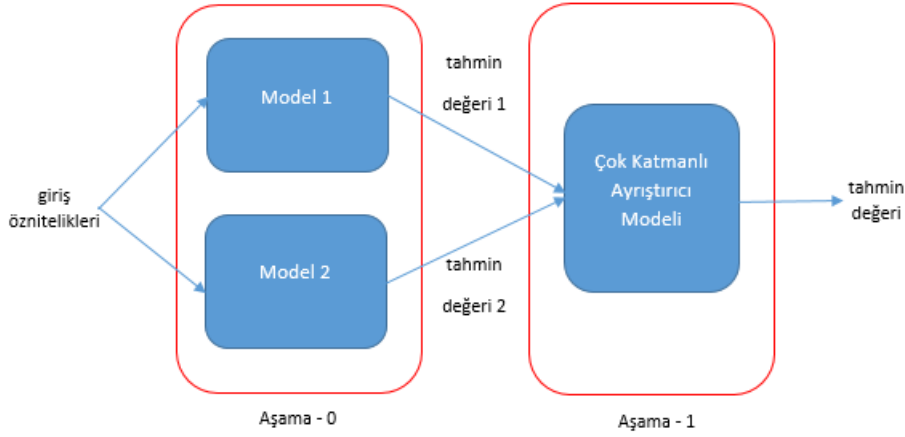


Şekil 4.3: Çalışmada uygulanan yığılmış genelleme.

Bir sonraki aşamada, aşama-0’da kullanılan dört algorithmadan bazılarının diğerlerine göre çoğunlukla daha yüksek başarımlar sağlaması gibi bir durumun var olup olmadığını belirlemek için bir deneme yapılmıştır. Bunun için öncelikle bu dört algoritmanın ikili kombinasyonları belirlenmiştir:

- Çok Katmanlı Ayrıştırıcı ve Bayes Ağı
- Çok Katmanlı Ayrıştırıcı ve Lineer Regresyon
- Çok Katmanlı Ayrıştırıcı ve Ardışık Minimal Optimizasyon
- Bayes Ağı ve Lineer Regresyon
- Bayes Ağı ve Ardışık Minimal Optimizasyon
- Ardışık Minimal Optimizasyon ve Lineer Regresyon

Bu altı farklı kombinasyon kullanılarak altı farklı deneme yapılmıştır. Her bir denemede ikili kombinasyonlar kümesinden bir eleman seçilmektedir. Seçilen eleman hangi iki algoritmayı içeriyorsa yığılmış genellemenin aşama-0’ında sadece o iki algoritma kullanılarak oluşturulmuş modeller yer almaktadır. Bu şekilde oluşturulan yığılmış genellemenin şeması Şekil 4.4’te görülmektedir.



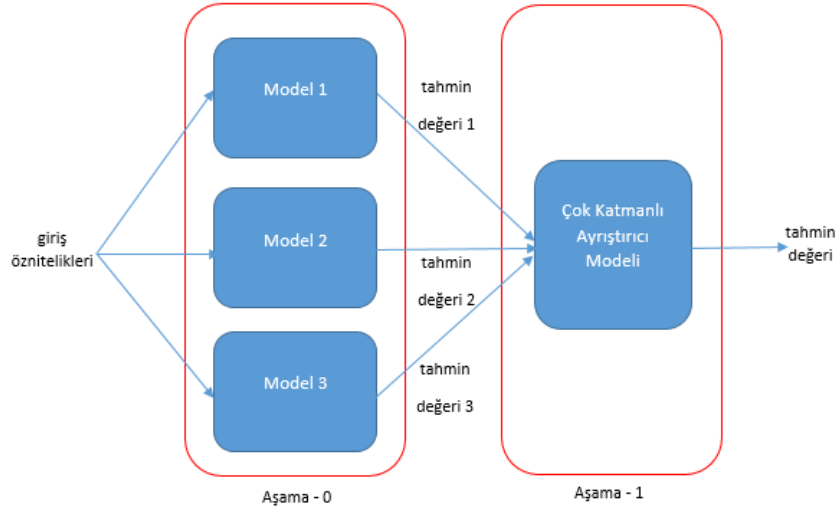
Şekil 4.4: İkili kombinasyonlarla yığılmış genelleme.

Ardından aynı yaklaşım dört algoritmanın üçlü kombinasyonları kullanılarak denenmiştir. Öncelikle dört algoritmanın üçlü kombinasyonları belirlenmiştir:

- Çok Katmanlı Ayırtıcı, Bayes Ağı, Ardışık Minimal Optimizasyon
- Çok Katmanlı Ayırtıcı, Bayes Ağı, Lineer Regresyon
- Çok Katmanlı Ayırtıcı, Ardışık Minimal Optimizasyon, Lineer Regresyon
- Bayes Ağı, Ardışık Minimal Optimizasyon, Lineer Regresyon

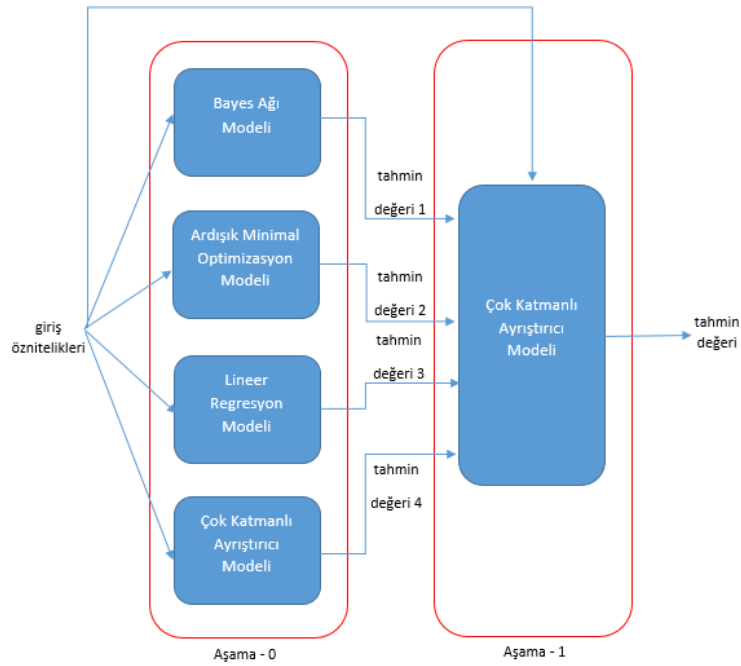
Bu denemede aşama-0'da sadece bu kümeden seçilen bir elemanın içerdiği üç farklı algoritma kullanılarak oluşturulan modeller yer almaktadır. Aşama-1'in girişine bu modellerden elde edilen tahmin sonuçları verilmektedir. İlgili yapı Şekil 4.5'te görülmektedir.

Son denemede ise yığılmış genellenenin farklılaştırılmış bir versiyonu olan ve bu çalışmada önerilen yeni yönteme ilişkin bir deneme yapılmıştır. Bu aşama yığılmış genellemede olduğu gibi aşama-0 ve aşama-1 olmak üzere iki aşama içermektedir. Bu yöntemdeki aşama-0, yığılmış genellemedeki aşama-0'ın tamamen aynıdır. Yöntem aşama-1'in girişi için özellik seçimi kısmında yığılmış genellemeden farklılaşmaktadır. Yığılmış genellemede, aşama-1'deki genelleme modelinin girişinde yalnızca aşama-0'daki modellerden elde edilen tahmin sonuçları kullanılmaktadır. Örneğin aşama-0'da dört farklı model olması durumunda aşama-1'in girişine yalnızca bu dört modelden elde edilen tahmin sonuçları verilmektedir. Bu çalışmada önerilen yeni yaklaşımda ise aşama-1'deki modelin girişinde sadece bu dört tahmin değerini kullanmak yerine aşama-0 modellerinin girişinde kullanılan özelliklerin de kullanılması yaklaşımı uygulanmaktadır.



Şekil 4.5: Üçlü kombinasyonlarla yığılmış genelleme.

Şekil 4.6’da görüleceği üzere aşama-0 modelinin girişinde x adet özellik bulunmaktadır. Aşama-0’da y adet farklı model bulunması durumunda y adet tahmin sonucu üretilecek ve aşama-1 modelinin girişine $x + y$ adet özellik verilecektir. Aşama-1’deki modelden elde edilen sonuç, tüm yöntemin nihai tahmin sonucunu verecektir.



Şekil 4.6: Çalışmada önerilen yöntem.

5. DENEYLER VE SONUÇLAR

Bu bölümde kullanılan veri kümesi hakkında bilgi verilecek ve yapılan denemeler ve bu denemelerin sonuçları incelenmektedir. Ayrıca bu çalışmada önerilen model topluluğu oluşturma (model birleştirme) yaklaşımı, yığılmış genelleme ve uygulanan diğer yöntemlerle karşılaştırılmaktadır.

5.1 Veri Kümesi

Bu çalışmada modellerin eğitimi ve testi için kullanılan veri kümesi, Türkiye’de hizmet veren bir ulusal kuruyemiş firmasının gerçek satış faturaları kullanılarak oluşturulmuştur. Bu veri kümesinin oluşturulmasında 2011, 2012 ve 2013 yıllarına ilişkin gerçek satış faturaları kullanılmıştır. Bu firma, Türkiye’nin hemen hemen her iline hizmet vermekte ve bünyesinde 98 adet ana dağıtım deposu bulundurmaktadır. Ayrıca bu firmanın 70 farklı ürünü vardır. Veri kümesi canlı veri tabanından alınmış toplamda 15317141 satır satış faturası kaydı ve 93987868 satır satış faturası detay kaydı içermektedir. Ana depo için seçilen özellik sayısı 12’dir. Bu özellikler ana deponun müşteri sayısı, ana deponun personel sayısı, araç sayısı, ana deponun müşteri yaşam boyu değeri, deponun müşterilerinin ortalama ağırlık değeri, deponun bulunduğu bölgedeki nüfus sayısı, o bölgedeki diğer ana depoların sayısı, ana deponun aylık ortalama satış miktarı (TL cinsinden), ana deponun sınıfı, deponun kaç metrekare satış alanına sahip olduğu, kaç adet alt depoya sahip olduğu ve deponun bulunduğu bölgedir.

5.2 Deneysel Sonuçlar

Deneylere başlamadan önce, modelleri eğitmek ve test etmek amacıyla veri kümesi 3 parçaya ayrılmıştır. Burada veri kümesi zamana göre sıralı şekilde ayrılmıştır. İlk veri kümesi 2011, ikinci veri kümesi 2012 ve üçüncü veri kümesi ise 2013 yılı verilerinden oluşmaktadır. Çalışmada kullanılan model yöntemleri ardışık iki makine öğrenmesi aşaması içermektedir ve bu ardışık modellerin aynı eğitim kümesiyle eğitilmesi doğru değildir. Bu nedenle oluşturulan 3 adet veri kümesinden ilki model

birleştirme yöntemlerinin aşama-0 modellerini eğitmek için kullanılırken, ikinci veri kümesi ise aşama-1'in eğitimi için kullanılmaktadır. Üçüncü veri kümesi ise birleştirilmiş modellerin testinde kullanılmaktadır. Çalışmada, birleştirilmiş modellerden önce bağımsız modellere ilişkin denemeler yapılmıştır. Bu denemeler yapılırken hem birinci hem ikinci veri kümeleri kullanılarak modeller eğitilmiş, üçüncü veri kümesi ile ise testler yapılmıştır.

Uygulanan yöntemlerin hata oranları MAPE (Ortalama Mutlak Yüzde Hata) ile hesaplanmıştır. MAPE hesaplamasında kullanılan formül denklem 5.1'de görülmektedir. Bu formülde A_t gerçek değerleri ifade ederken F_t ise tahmin sonucunda bulunan değerlere karşılık gelmektedir. Depo grupları için ayrı ayrı modeller oluşturularak yapılan testler için ilgili varyans değerleri de belirtilmiştir.

$$MAPE = \frac{100}{n} \sum_{t=1}^n \frac{|A_t - F_t|}{|A_t|} \quad (5.1)$$

Yöntem kısmında anlatıldığı üzere yapılan ilk denemelerde tüm depoların birlikte kullanıldığı bir Bayes ağı oluşturulmuştur. Burada modeller oluşturulurken oluşturulurken model giriş özellikleri olarak depo özellikleri, zaman bilgisi, ürün bilgisi kullanılmıştır. Modelin çıkışında ise gerçek satış miktarları kullanılmıştır. Ardından iki parçalı çizge demetleme yöntemiyle elde edilen ana dağıtım deposu gruplarının her biri için ayrı ayrı Bayes ağı modelleri oluşturulmuştur. Bir sonraki denemede ise alt dağıtım depoları grupları için bağımsız Bayes ağı modelleri oluşturularak testler yapılmıştır. Çizelge 5.1'de bu üç yaklaşımın hata oranları yer almaktadır.

Çizelge 5.1: Gruplar için ayrı modeller oluşturma ile tek model oluşturma ile kıyaslaması.

Yöntem	MAPE değeri
Tüm depolar için tek model	%49
Ana depolar için ayrı modeller	%24
Alt depolar için ayrı modeller	%17,5

Çizelgede görüldüğü üzere, en yüksek hata oranı tüm depoların birlikte kullanılmasıyla bir model oluşturulması durumunda elde edilmiştir. Ana depo ve sonrasında alt depolar için oluşturulan gruplar için farklı farklı modeller

oluşturulduğunda ise hata oranı düşmektedir. İlk denemede hata oranının yüksek olmasının farklı satış davranışlarına sahip depolar için tek bir model kullanılmasından kaynaklandığı düşünülmektedir. Bu nedenle ana dağıtım depoları satış davranışlarına göre gruplanıp bu gruplara özel modeller oluşturulduğunda hata oranında düşme gerçekleşmiştir. Bunun yanı sıra hata oranı hala yüksek olan depolar incelendiğinde bu durumun genellikle çok geniş bölgelere hizmet veren ve çok sayıda alt deposu bulunan depolarda görüldüğü farkedilmiştir. Bu problemi çözebilmek için alt depolar satış davranışlarına göre gruplandırılmış ve bu gruplar için ayrı modeller oluşturulması yöntemiyle başarımlar artırılmıştır.

Sonraki aşamada ise aynı yöntemin Bayes ağları yerine başka makine öğrenmesi algoritmaları kullanılarak gerçekleştirilmesi durumu incelenmiştir. Seçilen algoritmalar Ardışık Minimal Optimizasyon, Lineer Regresyon ve Çok Katmanlı Ayrıştırıcı'dır. Aynı yöntem bu algoritmalarla gerçekleştirildiğinde elde edilen sonuçlar Çizelge 5.2'de yer almaktadır. Çizelgede görüldüğü üzere genel hata oranı hesaplandığında en düşük hata oranının Bayes ağları ile elde edildiği görülmüştür.

Çizelge 5.2: 4 farklı modelin karşılaştırması.

Yöntem	MAPE değeri	Varyans değeri
Bayes Ağı	%17,5	0,0020
Çok Katmanlı Ayrıştırıcılar	%21,2	0,0034
Lineer Regresyon	%19,8	0,0029
Ardışık Minimal Optimizasyon	%23	0,0024

Yine yapılan incelemelerde, genelde hata oranının Bayes ağı ile daha düşük olduğu ancak bazı depolarda diğer algoritmalarla oluşturulmuş modellerin daha başarılı olduğu gözlenmiştir. Bu depolar arasında ortak bir özellik bulunup bulunmadığına yönelik çeşitli çalışmalar yapılmış ancak bir sonuç elde edilememiştir. Buradan yola çıkılarak farklı modellerin birlikte kullanılmasının genel başarımları arttırabileceği düşünülmüştür. Bu doğrultuda yığılmış genelleme algoritması kullanılarak bahsedilen dört algorithmadan oluşturulan modeller birleştirilmiştir. Bu çalışma sonucunda MAPE ile ölçülen hata oranı %21,9 olarak bulunmuştur. Bu test için varyans değeri 0,0025'dir.

Yine yapılan incelemelerde, genelde hata oranının Bayes ağı ile daha düşük olduğu Sonraki aşamada ise seçilen bu dört algoritmanın ikili kombinasyonları alınmış ve bu kombinasyonların her biri için ayrı ayrı yığılmış genelleme yöntemi uygulanmıştır.

Burada amaç tüm algoritmalarından oluşturulan modellerin birleştirilmesi yerine sadece bazı algoritma kombinasyonlarından oluşturulan modeller birleştirildiğinde başarımda nasıl bir değişim olacağını gözlemleyebilmektir. Ayrıca, bazı algoritmaların diğerlerine kıyasla her kombinasyonda kötü veya iyi sonuç vermesi gibi bir durum olup olmadığı belirlenmeye çalışılmıştır. İkili kombinasyonlar sonucunda elde edilen sonuçlar Çizelge 5.3'te görülmektedir.

Çizelge 5.3: İkili kombinasyonlarla yığılmış genelleme karşılaştırması.

Yöntem	MAPE değeri	Varyans değeri
{Bayes Ağı & Ardışık Minimal Optimizasyon}	%22	0,0026
{Bayes Ağı & Çok Katmanlı Ayrıştırıcı}	%20,7	0,0024
{Bayes Ağı & Lineer Regresyon}	%20,5	0,0026
{Ardışık Minimal Optimizasyon & Çok Katmanlı Ayrıştırıcı}	%23,8	0,0030
{Ardışık Minimal Optimizasyon & Lineer Regresyon}	%23	0,0031
{Çok Katmanlı Ayrıştırıcı & Lineer Regresyon}	%22,8	0,0027

Aynı çalışma, seçilen dört algoritmanın üçlü kombinasyonları kullanılarak da yapılmıştır. Bu kombinasyonlarla oluşturulan modellerden elde edilen hata oranları Çizelge 5.4'te görülmektedir.

Çizelge 5.4: Üçlü kombinasyonlarla yığılmış genelleme karşılaştırması.

Yöntem	MAPE değeri	Varyans değeri
{Bayes Ağı, Ardışık Minimal Optimizasyon, Çok Katmanlı Ayrıştırıcı}	%22,6	0,0029
{Bayes Ağı, Ardışık Minimal Optimizasyon, Lineer Regresyon}	%21,4	0,0026
{Bayes Ağı, Çok Katmanlı Ayrıştırıcı, Lineer Regresyon}	%20,3	0,0025
{Ardışık Minimal Optimizasyon, Çok Katmanlı Ayrıştırıcı, Lineer Regresyon}	%23,4	0,0031

Yapılan son denemede ise bu çalışmada önerilen birleştirme yöntemi kullanılmıştır. Daha önce bahsedildiği üzere bu yöntem ile aşama-1'deki genelleme modelinin girişinde aşama-0'daki modellerden elde edilen tahmin sonuçlarının yanı sıra aşama-0'ın girişinde kullanılan özellikler de kullanılmaktadır. Bu yöntemin diğer yöntemlerle kıyaslaması Çizelge 5.5'te yer almaktadır.

Çalışmada önerilen modelin diğerlerine kıyasla daha yüksek başarımlar sağladığı görülmektedir. Bu durumun, aşama-0 girişindeki özelliklerin ve aşama-0'dan elde edilen tahmin sonuçlarının aşama-1'de birlikte kullanılmasıyla aşama-1'in bir optimizasyon katmanı gibi çalışmasından kaynaklandığı düşünülmektedir. Yığılmış genellemede ise önerilen yöntemden farklı olarak, aşama-1'deki model aşama-0'dan elde edilen tahmin değerlerini gerçek tahmin değerlerine eşleme amacıyla kullanılmaktadır.

Çizelge 5.5: Önerilen modellerle karşılaştırma.

Yöntem	MAPE değeri	Varyans değeri
Önerilen Yöntem	%12,7	0,0002
Yığılmış Genelleme	%21,9	0,0025
Bayes Ağı	%17,5	0,0020
Bayes Ağı ve Lineer Regresyon ile Yığ. Gen.	%20,5	0,0026
Bayes Ağı, Çok Kat. Ayrış., Lineer Reg. ile Yığ. Gen.	%20,3	0,0025

6. SONUÇ VE ÖNERİLER

Tez kapsamında yapılan çalışmada amaç, ileri tarihli bir zaman dilimi için hangi ana deponun hangi üründen kaç adet satacağını öngörebilen bir model geliştirmektir. Talep tahmini, tedarik zinciri yönteminin oldukça zor ve karmaşık bir alt sürecidir. Bu problemin zorluklarından biri, depo için talep tahmini sırasında göz önünde bulundurulması gereken faktörlerin fazlalığıdır. Buna ek olarak çalışmada kullanılan ve Türkiye'deki ulusal bir kuruyemiş firmasının üç yıllık satış verilerinden oluşturulan veri kümesinde oldukça fazla sayıda ana dağıtım deposu ve ürün bulunması problemin zorluğunu arttırmaktadır.

Çalışmada talep tahmini için öncelikle ana dağıtım depoları ve alt dağıtım depoları iki parçalı çizge demetleme algoritması kullanılarak gruplanmıştır. Bu gruplar için ayrı ayrı modeller oluşturmanın tüm depolar için tek bir model oluşturmaya kıyasla daha yüksek başarımlar sağladığı görülmüştür. Çalışmanın sonraki aşamasında ise yığılmış genellemeden yola çıkılarak oluşturulmuş bir model birleştirme yöntemi önerilmiştir. Önerilen yöntem de yığılmış yöntemde olduğu gibi iki aşama içermektedir. Bu yöntemde yığılmış genellemeden farklı olarak aşama-1'deki genelleme yönteminin girişinde aşama-0'dan elde edilen tahmin sonuçlarına ek olarak aşama-0'ın girişinde kullanılan özellikler de kullanılmaktadır. Çalışmada daha önce ana depolar için talep tahmininde kullanılmamış olan yığılmış genelleme yöntemi de denenmiştir. Böylelikle depo talep tahmini için çalışmada önerilen yöntem ile yığılmış genelleme yöntemi karşılaştırılmıştır.

Deneysel sonuçlara göre çalışmada önerilen ana ve alt depoları satış davranışlarına göre iki parçalı çizge demetleme kullanarak gruplama ve bu gruplar için ayrı modeller oluşturma yaklaşımı başarımları önemli ölçüde arttırmıştır. Buna ek olarak, oluşturulan modelleri birleştirmek için kullanılan ve bu çalışmada önerilen yöntemin yığılmış genellemeye kıyasla daha düşük hata oranı sağladığı görülmüştür.

Konuyla ilgili olarak gelecek dönemde yapılacak çalışmalardan biri depoların müşterilerinin de çeşitli veri madenciliği yöntemlerinden yararlanılarak gruplandırılması (satış hacmi, hedef kitlesi, satış davranışı vb.) ve bu grupların da

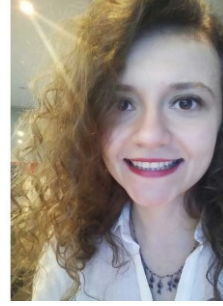
tahmin modeline dahil edilmesi olabilir. Buna ek olarak ürünlerin sınıflandırılmasında ilgili bölgedeki pazar payları ve karlılıkları göz önünde bulundurularak çeşitli veri madenciliği çalışmaları yapılabilir. Bu çalışmalar sonucunda oluşturulan farklı tahmin modellerinin de model topluluklarına dahil edilmesiyle başarımın nasıl değişeceğine ilişkin testler yapılabilir.

KAYNAKLAR

- [1] *Logistics and Supply Chain Management* (1992). London: Pitman Publishing.
- [2] **Mentzer, J. T., DeWitt W., Keebler, J. S., Min, S., Nix, N. W., Smith, C.D. & Zacharia, Z. G.** (2001). Defining Supply Chain Management, *Journal of Business Logistics*, 22(2).
- [3] **Liu, L. M., Bhattacharyya S., Sclove, S. L., Chen, R. & Lattyak, W.J.** (2001). Data Mining on Time Series: An Illustration Using Fast-Food Restaurant Franchise Data, *Computational Statistics & Data Analysis*, 37, 455-476.
- [4] **Ali, Ö. G., Sayın, S., Woensel, T.V. & Fransoo, J.** (2009). SKU Demand Forecasting in the Presence of Promotions, *Expert Systems with Applications*, 36(10), 12340-12348.
- [5] **Johnson, K., Lee, B. H. A. & Simchi-Levi, D.** (2014). Analytics for an Online Retailer: Demand Forecasting and Price Optimization.
- [6] **Chang, P. C., Wang Y. W. & Liu, C. H.** (2007). The Development of a Weighted Evolving Fuzzy Neural Network for PCB Sales Forecasting, *Expert Systems with Applications*, 32, 86-96.
- [7] **Sun, Z. L., Choi T. M. & Yu, Y.** (2008). Sales Forecasting Using Extreme Learning Machine With Applications In Fashion Retailing, *Decision Support Systems*, 46, 411-419.
- [8] **Yu, Y., Choi T. & Hui, C.** (2011). An Intelligent Fast Sales Forecasting Model for Fashion Products, *Expert Systems with Applications*, 38, 7373-7379.
- [9] **Ling, S. H.** (2010). *Genetic Algorithm and Variable Neural Networks: Theory and Application*. Lambert Academy Publishing.
- [10] **Au, K. F., Choi T. M. & Yu, Y.** (2008). Fashion Retail Forecasting by Evolutionary Neural Networks, *International Journal of Production Economics*, 114, 615-630.
- [11] **Gutierrez, R. S., Solis, A. & Mukhopadhyay, S.** (2008). Lumpy Demand Forecasting Using Neural Networks, *International Journal of Production Economics*, 111, 409-420.
- [12] **Hasin, M. A. A., Ghosh, S. & Shareef, M.A.** (2011). An ANN Approach to Demand Forecasting in Retail Trade in Bangladesh, *International Journal of Trade, Economics and Finance*, 2(2).
- [13] **Efendigil, T., Önüt, S. & Kahraman, C.** (2009). A Decision Support System for Demand Forecasting with Artificial Neural Networks and Neuro-Fuzzy Models: A Comparative Analysis, *Expert Systems with Applications*, 36(3), 6697-6707.

- [14] **Joseph, A., Singh E., & Larrain, M.** (2007). Forecasting Aggregate Sales with Interest Rates Using Multiple Neural Networks, *Intelligent Engineering Systems Through Artificial Neural Networks*, 17.
- [15] **Arunraj, N. S. & Ahrens D.** (2015). A Hybrid Seasonal Autoregressive Integrated Moving Average and Quantile Regression for Daily Food Sales Forecasting, *International Journal of Production Economics*, 170, 321-335.
- [16] **Jaipuria, S. & Mahapatra S. S.** (2014). An Improved Demand Forecasting Method to Reduce Bullwhip Effect in Supply Chains, *Expert System with Applications*, 41(5), 2395-2408.
- [17] **Doganis, P., Alexandridis, A., Patrinos, P. & Sarimveis, H.** (2006). Time Series Sales Forecasting For Short Shelf-Life Food Products Based On Artificial Neural Networks And Evolutionary Computing, *Journal Of Food Engineering*, 75, 196-204.
- [18] **Aburto, L., Weber, R.** (2007). Improved Supply Chain Management Based On Hybrid Demand Forecasts, *Applied Soft Computing*, 7(1), 136-144.
- [19] **Dietterich, T. G.** (2000). Ensemble Methods in Machine Learning. *Multiple Classifier Systems*, (pp.1-15), Springer Berlin Heidelberg.
- [20] **Roli, F., Giacinto, G. & Vernazza, G.** (2001). Methods for Designing Multiple Classifier Systems. *Multiple Classifier Systems*, (pp.78-87), Springer Berlin Heidelberg.
- [21] **Kuncheva, L. I.** (2004). *Combining Pattern Classifiers: Methods and Algorithms*. Wiley & Sons.
- [22] **Breiman, L.** (2001). Random Forests, *Machine Learning*, 45(1), 5-32.
- [23] **Caruana, R., Niculescu-Mizil, A., Crew, G. & Ksikes, A.** (2004). Ensemble Selection from Libraries of Models, *ICML '04*, (pp.18), NewYork.
- [24] **Wolpert, D. H.** (1992). Stacked Generalization, *Neural Networks*, 5, 241-259.
- [25] **Gama, J. & Brazdil, P.** (1999). Linear Tree, *Intelligent Data Analytics*, 3(1), 1-22.
- [26] **Dzeroski, S. & Zenko, B.** (2004). Is Combining Classifiers Better Than Selecting The Best One?, *Machine Learning*, 54(3), 195-209.
- [27] **Ghorbani, A. & Owrangh, K.** (2001). Stacked Generalization In Neural Networks: Generalization On Statistically Neutral Problems, *IJCNN '01. International Joint Conference on Neural Networks*, (pp.1715-1720). IEEE, September 17-19.
- [28] **Ting, K. M. & Witten, I. A.** (1999). Issues in Stacked Generalization, *Journal Of Artificial Intelligence Research*, 10, 271-289.
- [29] **Carroll, M. K. & Cha, S. H.** (2003). Application of Stacked Generalization to a Protein Localization Prediction Task, *Proceedings of the 7th JCIS*, (pp.13-16).

- [30] **Gandhi, H., Green, D., Kounios, J., Clark, C. M. & Polikar, R.** (2006). Stacked Generalization for Early Diagnosis of Alzheimer's Disease, *Proceedings of the 28th IEEE EMBS Annual International Conference*, Aug 30-September 3.
- [31] **Kotsiantis, S., Koumanakos, E., Tezlepis, D. & Tampakas, V.** (2006). Forecasting Fraudulent Financial Statements Using Data Mining, *International Journal of Computational Science*, 3(2), 104-110.
- [32] **Tsai, C. F.** (2005). Training Support Vector Machines Based On Stacked Generalization For Image Classification, *Neurocomputing*, 64, 479-503.
- [33] **Doumpos, M. & Zopounidis, C.** (2007). Model Combination For Credit Risk Assessment: A Stacked Generalization Approach, *Annals of Operations Research*, 151(1), 289-306.
- [34] **Sakkis, G., Androutsopoulos, I., Paliouras, G., Karkaletsis, V., Spyropoulos, C. & Stamatopoulos, P.** (2001). Stacking Classifiers for Anti-Spam Filtering of E-mail, *Empirical Methods in Natural Language Processing*, (pp.44-50).
- [35] **Zha, H., He, X., Ding, C., Simon, H. & Gu, M.** (2001). Bipartite Graph Partitioning and Data Clustering, *CIKM'01*, (pp.25-32), November 5-10.
- [36] **Friedman, N., Geiger, D. & Goldszmidt, M.** (1997). Bayesian Network Classifiers, *Machine Learning*, 29(2-3), 131-163.
- [37] **Haykin, S.** (1998). *Neural Networks: A Comprehensive Foundation*. Prentice Hall.
- [38] **Freedman, D. A.** (2009). *Statistical Models: Theory and Practice*. Cambridge University Press.
- [39] **Platt, John.** (1999). Fast Training Of Support Vector Machines Using Sequential Minimal Optimization, *Advances in kernel methods-support vector learning*, 3.
- [40] **Url-1** <<http://protege.stanford.edu>>, erişim tarihi 26.11.2015.



ÖZGEÇMİŞ

Ad-Soyad : İrem İşlek
Doğum Tarihi ve Yeri : 21.02.1990 Edirne
E-posta : isleki@itu.edu.tr

ÖĞRENİM DURUMU:

- **Lisans** : 2013, İstanbul Teknik Üniversitesi, Bilgisayar Bilişim Fakültesi, Bilgisayar Mühendisliği

MESLEKİ DENEYİM VE ÖDÜLLER:

- 2013 yılında lisans bitirme çalışmasıyla İdea Teknoloji Çözümleri Bitirme Tezi Destek Programı kapsamında ödül aldı.
- 2013 yılından beri İdea Teknoloji Çözümleri'nde Ar-Ge Mühendisi olarak çalışmaktadır.
- 0484.STZ.2013-2 numaralı SANTEZ projesinde tez öğrencisi olarak yer aldı.

YÜKSEK LİSANS TEZİNDEN TÜRETİLEN YAYINLAR, SUNUMLAR VE PATENTLER:

- İşlek, İ., Öğüdücü, Ş., 2015. A Retail Demand Forecasting Model Based On Data Mining Techniques, *The 24th IEEE International Symposium on Industrial Electronics*.