Journal of the Faculty of Engineering and Architecture of Gazi University 39:3 (2024) 1365-1374

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

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

A data transformation method for multi-perspective process mining healthcare applications

Tuğba Gürgen Erdoğan*
Department of Computer Engineering, Faculty of Engineering, Hacettepe University, 06800, Ankara, Türkiye

Highlights:

- Multi-perspective process mining application
- Data transformation method
- Healthcare Process Analysis

Keywords:

- Process Mining
- Process Discovery
- Healthcare data modelling
- Data Privacy
- Performance Analysis

Article Info:

Research Article
Received: 12.08.2022
Accepted: 23.07.2023

DOI:

10.17341/gazimmfd.1161239

Correspondence:

Author: Tuğba Gürgen Erdoğan
e-mail: tugba@cs.hacettepe.edu.tr
phone: +90 533 200 8572

Graphical/Tabular Abstract

The applications of process mining, which is a business process management technique, in the field of healthcare are increasing day by day. In process mining, it is possible to analyze the process for three main purposes: discovery of the process, compliance control and process improvement, based on event logs recorded in information systems. Transforming patient-based healthcare processes data into process and event-based event logs is the first step of a process mining project in order to apply process mining techniques to human-centered, distributed, complex and multidisciplinary healthcare processes and to improve the quality of health care services. The process model discovered in multi-perspective process mining is expanded from different perspectives such as control flow, organizational, data, time and function, and the discovered process becomes more understandable. In this study, and in order to apply multi-perspective process mining, a method is proposed to convert the healthcare processes data recorded in hospital information systems into event logs and the steps of the method and outputs are shown in Figure A.

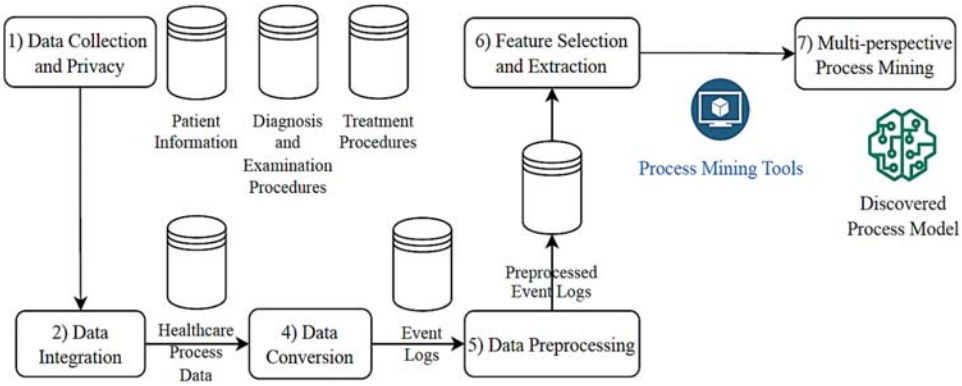


Figure A. Data transformation method overview

Purpose: The purpose of the data transformation method proposed in this study is to explain the work steps required to convert health process data into event logs to apply multi-perspective process mining techniques to healthcare processes.

Theory and Methods: The data transformation method consists of six steps: data collection and data privacy, data integration, data conversion, data preprocessing, feature selection and extraction, and multi-perspective process mining analysis. The proposed method is validated by a case study by transforming the surgery process data of a university hospital in Turkey into event logs. Process discovery, process variation analysis, and performance analysis techniques are applied to the surgery process data. The actual process is discovered, and the applicability of the data transformation method is shown on the real data.

Results: Healthcare data of the surgery process are collected, and data privacy issues are addressed and the data are anonymized. Healthcare data which are anonymized and collected from four different sources, are combined in a patient-oriented manner. Activity names are determined by using domain knowledge during data conversion. Since the converted data are multidimensional, 14 features are selected among 29 features and the complexity of the data is reduced. These features are selected together with the domain expert based on the hospital's process definitions and measured key performance indicators. The discovered process model is illustrated by combining two different perspectives.

Conclusion: The proposed method is developed specifically for health data and enriched with data security issues, data science methodology, and data science techniques. The method includes guiding approaches by including a practical example to address the inherent complexity, distributed, and multidisciplinary nature of health data. All steps required to convert object-based surgery process data into event logs are explained in detail and these processes are illustrated with a practical example.

Journal of the Faculty of Engineering and Architecture of Gazi University 39:3 (2024) 1365-1374

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

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

Çok perspektifli süreç madenciliği sağlık uygulamaları için bir veri dönüştürme yöntemi

Tuğba Gürgen Erdoğan*
Hacettepe Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, 06800, Çankaya, Ankara, Türkiye

Ö N E Ç I K A N L A R

- Çok perspektifli süreç madenciliği
- Veri dönüşümü
- Sağlık Süreci

Makale Bilgileri
Araştırma Makalesi
Geliş: 12.08.2022
Kabul: 23.07.2023

DOI:
10.17341/gazimmfd.1161239

Anahtar Kelimeler:
Süreç madenciliği,
sağlık süreçleri,
süreç keşfi,
veri güvenliği

ÖZ

Bir iş süreci yönetim tekniği olan süreç madenciliğinin sağlık alanında uygulamaları her geçen gün artmaktadır. Süreç madenciliğinde, bilgi sistemlerinde kaydedilen olay günlüklerinden hareketle sürecin keşfedilmesi, uygunluk kontrolü ve süreç iyileştirme olmak üzere üç temel amaçla, süreci analiz etmek mümkün olmaktadır. İnsan odaklı, dağıtık, karmaşık ve çok disiplinli sağlık süreçleri verisine süreç madenciliği tekniklerini uygulamak ve sağlık hizmetlerinin kalitesini arttırmak için hasta tabanlı sağlık süreçleri verisinin süreç ve olay tabanlı olay günlüğüne dönüştürülmesi, bir süreç madenciliği projesinin ilk adımıdır. Çok perspektifli süreç madenciliğinde keşfedilen süreç modeli kontrol akışı, örgütsel, veri, zaman ve fonksiyon gibi farklı perspektiflerden genişletilerek, keşfedilen süreç daha anlamlı hale gelmektedir. Bu çalışmada çok perspektifli süreç madenciliği uygulamak adına hastane bilgi sistemlerinde dağıtık olarak kaydedilen sağlık süreçleri verisini olay günlüğüne dönüştürme için bir yöntem önerilmiştir. Veri dönüştürme yöntemi; veri toplama ve veri güvenliği, verinin bütünleştirilmesi, veri dönüştürme, veri ön işleme, özellik seçimi ve çıkarımı ve çok perspektifli süreç madenciliği analizi olmak üzere altı adımdan oluşmaktadır. Önerilen yöntemin doğrulaması, Türkiye'deki bir üniversite hastanesinde yapılan durum çalışması ile yapılmıştır. Durum çalışmasına ait ameliyat süreci verisi üzerinde süreç keşfi algoritmaları uygulanmış ve veri dönüştürme yönteminin uygulanabilirliği gerçek veri üzerinde gösterilmiştir. Yöntemin, çok perspektifli süreç madenciliğinin Türkiye'de sağlık alanındaki uygulamalarına katkıda bulunması beklenmektedir.

A data transformation method for multi-perspective process mining healthcare applications

H I G H L I G H T S

- Multi-perspective process mining
- Data transformation
- Healthcare Process

Article Info
Research Article
Received: 12.08.2022
Accepted: 23.07.2023

DOI:
10.17341/gazimmfd.1161239

Keywords:
Process mining,
healthcare process,
process discovery,
data privacy

ABSTRACT

The applications of process mining, which is a business process management technique, in the field of healthcare are increasing day by day. In process mining, it is possible to analyze the process for three main purposes: discovery of the process, compliance control and process improvement, based on event logs recorded in information systems. Transforming patient-based healthcare processes data into process and event-based event logs is the first step of a process mining project in order to apply process mining techniques to human-centered, distributed, complex and multidisciplinary healthcare processes and to improve the quality of health care services. The process model discovered in multi-perspective process mining is expanded from different perspectives such as control flow, organizational, data, time and function, and the discovered process becomes more understandable. In this study, and in order to apply multi-perspective process mining, a method is proposed to convert the healthcare processes data recorded in hospital information systems into event logs. Data transformation method consists of six steps: data collection and data privacy, data integration, data conversion, data preprocessing, feature selection and extraction, and multi-perspective process mining analysis. The proposed method was validated by a case study by transforming the surgery process data of a university hospital in Turkey into an event log. The process discovery algorithm was applied to the surgery process data of the case study and the actual process was discovered, and the applicability of the data transformation method was shown on the real data. With the guiding feature of the method for healthcare professionals, it is expected to contribute to the applications of multi-perspective process mining in the field of healthcare in Turkey.

*Sorumlu Yazar/Yazarlar / Corresponding Author/Authors : *tugba@cs.hacettepe.edu.tr / Tel: +90 533 200 8572

1. Giriş (Introduction)

Süreç madenciliği makine öğrenmesi ve veri madenciliği yöntemleri ile beslenen bir süreç yönetim tekniğidir [1]. Bilgi sistemlerinde kaydedilen veri, süreç madenciliği tekniklerinin en önemli kaynağıdır. Süreç madenciliği tekniklerini uygulamak için verinin olay günlüğü olarak temsil edilmesi gerekmektedir. Olay günlüklerinden hareketle süreç madenciliği çalışmalarının sürecin keşfi, uygunluk kontrolü ve sürecin iyileştirilmesi olmak üzere üç temel amacı vardır [2]. Birçok alanda kullanımı artan süreç madenciliği tekniğinin sağlık alanında da yaygın olarak kullanıldığı birçok literatür taraması çalışmasıyla gösterilmiştir [3-6].

Her olay kaydı bir etkinliği temsil eder ve her etkinlik bir sürecin parçasıdır. Olay günlükleri, sürece ilişkin olaylar hakkında olayları başlatan ve gerçekleştiren kişi veya kuruluş, olayın başlama ve bitiş zamanı ve yaş, cinsiyet, büyüklük gibi veri unsurları gibi ayrıntılı bilgileri içermektedir. Olay günlüklerindeki tüm bu bilgiler, süreçleri farklı perspektiflerden keşfetmek için kullanılmaktadır. Süreç madenciliğinde kontrol akışı, örgütsel, veri, zaman, uygunluk ve fonksiyon gibi birçok perspektif vardır [7]. Her bir perspektif için ayrı ayrı madencilik sonuçları sağlayan çeşitli keşif algoritmaları bulunmaktadır. Bununla birlikte, perspektifler izole değildir ve hepsi birbiriyle ilişkilidir. Perspektifler kendilerine özgü olarak farklı türde veriler gerektirdiğinden, üretilen çıktılar da çeşitli modelleme dilleri ve notasyonları ile gösterilebilir. Bunlardan bazıları Petri ağı [8], Sezgisel ağ [9], Bulanık model [10], BPMN [11] ve Data Petri ağı [12] modelleme gösterimleridir. Bütünsel bir süreç modeli elde etmek, diğer bir deyişle birden çok perspektifi içeren bir model oluşturmak, için öncelikle süreç keşfi gerçekleştirilmelidir. Sık kullanılan süreç madenciliği tekniklerinden Inductive Miner [13] ya da Alpha Miner [14] ile gerçekleşen sürece ait kontrol-akış perspektifini resmeden bir süreç modeli oluşturulur. Bu model, sürecin iskelet modeli olarak adlandırılır ve iskelet modele veri, kaynak ve performans gibi diğer perspektifler eklenerek keşfedilen süreç modeli genişletilir. Genişletilen süreç modelinde, süreci ait sadece olayların akış sırası ve yolları izleyen durumların sayısı gibi kontrol-akışı perspektifine ek olarak; etkinlik boyunca beklenen süreler, etkinliği gerçekleştiren roller, etkinlikler ile ilişkili veriye ait bilgiler de resmedilmektedir. Tüm bu bilgiler ışığında süreci daha iyi anlamak mümkün olmaktadır.

Çok sayıda alanda yaygın olarak kullanılan süreç madenciliğinin sağlık alanında uygulanması, 2008 yılında Mans vd. [15] bir Alman hastanesinde süreç madenciliği uygulaması ile dikkat çekmeye başlamış, günümüze kadar da artarak devam etmiştir [5], [16-19]. Süreç madenciliğinin sağlık alanındaki uygulamalarında ve ikinci çalışmalarda belirtilen en temel zorluklar ise sağlık süreci verisinin özellikleri ile ilişkilendirilmiştir [4]. Yapısal olmayan, dağıtık, karmaşık ve çok disiplinli sağlık verisini [20] süreç madenciliği tekniklerine uygulamak için hastane bilgi sistemlerinde kaydedilen nesne odaklı verinin olay tabanlı veriye dönüştürülmesi, bir süreç madenciliği projesinin en çok emek isteyen ve zaman alan kısmı olarak tanımlanabilir [19, 21]. Sağlık uygulamalarında, bir süreç madenciliği projesinin ötesine geçerek çok perspektifli süreç madenciliği tekniklerini uygulamak ve sağlık süreci modelini veri, zaman gibi farklı perspektifler ile zenginleştirmek, keşfedilen süreci daha anlamlı hale getirmektedir [22-24]. Bu çalışmada deneyimlerimize ve literatüre dayanarak hastane bilgi sisteminden elde edilen nesne tabanlı sağlık süreçleri verisinin olay tabanlı veriye dönüştürme adımlarını içeren bir yöntem önerilmiştir. Önerilen yöntem adımları, durum çalışması tasarım yaklaşımı [25, 26] izlenerek ameliyat verisi sürecine uygulanmış ve önerilen yöntem doğrulanmıştır. Keşfedilen ameliyat süreci verisi kontrol akış ve zaman perspektifleri ile resmedilmiştir. Bu makalenin geri kalanı şu şekilde düzenlenmiştir: 2. bölüm, süreç madenciliği, olay ve sağlık verileri hakkında ön bilgiler sunmakta ve sağlık süreçleri alanında süreç madenciliği uygulamaları için önerilmiş veri dönüştürme yöntemlerine genel bir bakış sağlamaktadır. 3. bölüm, çok perspektifli süreç madenciliği tabanlı veri dönüştürme yönteminin adımlarını ayrıntılı olarak açıklamaktadır. 4. bölüm, ameliyat süreci ve sonuçları üzerine bir durum çalışması ile pratik bir örnek içermektedir. Son bölüm, gelecek araştırma fırsatlarını belirlemek için durum çalışmasının sonuçlarını tartışmakta ve gelecek çalışmalar için genel sonuçları ve önerileri sunmaktadır.

2. Ön Bilgi ve İlişkili Çalışmalar (Background and Related Work)

2.1. Ön Bilgi (Background)

Süreç madenciliği, bilgi sistemlerinde kaydedilen veriden hareketle sürecin keşfedilmesi, uygunsuzların tespit edilmesi ve sürecin iyileştirilmesi amaçları ile uygulanan bir veri bilimi tekniğidir. Bir yanda diğer veri bilimi teknikleri arasından yapay zekâ ve veri madenciliği, diğer yanda süreç modelleme ve analizi arasında yer alan nispeten yeni bir araştırma alanıdır [1]. Yapay zekâ ve veri madenciliğinin aksine veri süreç odaklı işlenmektedir. Süreç modelleme ve analizi tekniklerinden farklı olarak ise akan süreç verisi analize dâhil edilir

Süreç madenciliği, bir olay günlüğünden süreç modellerinin çıkarılması, model ve olay günlüğünü karşılaştırarak sapmaların izlenmesi, sosyal ağ/organizasyonel madenciliği, simülasyon modellerinin otomatik inşası gibi teknikleri içeren bir süreç yönetim tekniğidir. ProM [27], Disco [28], PM4Py [29], Celonis [30] ve RBupar [31] gibi çok sayıda açık kaynaklı veya ticari süreç madenciliği araçları ile süreç madenciliği algoritmalarını uygulamak mümkündür. Temel olarak süreç madenciliğinin süreç keşfi, uygunluk kontrolü ve süreç iyileştirme olmak üzere üç temel amacı vardır: Ayrıca süreç madenciliği, çevrimiçi veya çevrimdışı olarak da uygulanmaktadır. Süreç madenciliğinin sonuçları, süreçler hakkında akıl yürütmek, yeniden tasarlamak veya operasyonel destek için kararlar almak için kullanılmaktadır. En çok tercih edilen süreç yollarını ve darboğazları teşhis ederek, gecikmeleri ortaya çıkararak, çakışmaları veya tekrarları belirleyerek, keşfedilen süreç haritalarını gözden geçirerek ve tekrarlı çalışmayı ortadan kaldırarak süreçlerin iyileştirilmesi, süreç madenciliğinin organizasyonlara sağladığı faydalar arasındadır.

Süreç madenciliğini uygulamanın bir başka amacı da süreçlerdeki başarım veya uyum sorunlarının kök nedenlerini bulmaktır [32]. Süreç madenciliği ile kök-neden analizi, mevcut bir olay günlüğü ile başlar ve olay günlüğünü ilgili bağlamsal bilgilerle zenginleştirir. Süreç madenciliği, kontrol akışı, örgütsel, veri, zaman ve fonksiyona odaklanan çeşitli perspektiflerden uygulanmaktadır [7, 33, 34]. Bir süreç modelini birden fazla perspektifle keşfetmek ve resmetmek, çok perspektifli süreç madenciliği olarak adlandırılmaktadır. Kontrol akışı perspektifi, olayların sırasına ve süreç modelinin keşfine odaklanır ve olası tüm yollar için en iyi model tanımlarını bulmayı hedefler. Örgütsel perspektif, süreç modelini rollere ve organizasyona göre sınıflandırmak için aktörler (insan, sistem ve rol) ve bu aktörler arasındaki ilişkilerle ilgilidir. Veri perspektifi, vakaların tanımına ve gerçek verileri etkileyen faktörlere odaklanır. Zaman perspektifi, olayların meydana gelme zamanı ve sıklıkları ile ilgilidir. Genellikle darboğazları belirlemeye, hizmet seviyelerini ölçmeye, kaynakların kullanımını izlemeye ve devam eden olaylar için kalan süreyi tahmin etmeye yardımcı olur.

Tüm süreç madenciliği perspektiflerini bir araya getirmek için, keşif adımındaki olay günlüğünden tüm olası aktivite yollarını bulan kontrol akışı perspektifiyle akış başlatılır. Daha sonra oluşturulan

süreç modelinin uygunluk kontrolü adımında gerçekliğin modele uygun olup olmadığı kontrol edilir ve varsa sapmalar tespit edilir. Diğer bir yandan ise, iyileştirme/geliştirme adımında süreç modeline ek perspektifler uygulanır.

Bilgi sistemlerinde nesnelere ilişkin süreçlere ait bilgiler *olay verileri* olarak adlandırılır [35]. Bir yazılım konfigürasyon yönetimi aracında nesne yazılımcı ya da görev bilgisi iken, hastane bilgi sisteminde nesne hasta ve hastanın sistemde geçirmiş olduğu etkinliklerdir. Süreç madenciliğinin bilgi sisteminde kaydedilen olay verisine uygulanabilmesi için olay verisinin *olay günlüğüne* dönüştürülmesi gerekir. Her olay bir etkinliği, her etkinlik ise bir sürecin bir parçasını temsil etmektedir. Olay günlükleri, olayların kaynağı; yani etkinliği başlatan ve gerçekleştiren kişi ya da araç ve türü, boyutu ya da yorumları gibi veri nitelikleri hakkında ayrıntılı bilgi tutmaktadır [35]. Olaya ait zaman bilgisi, olay günlüklerinin kritik özellikleridir ve olayların sırası gerçekleşen sürecin iskeletini oluşturur. Olay günlüğü ayrıca olaya ait diğer veri, kaynak, rol ve kurum bilgilerini ve performans bilgilerini içermektedir. Zaman ve etkinlik bilgileri sürecin kontrol akışını gösterirken, diğer bilgiler süreci durum, kaynak/rol gibi farklı perspektiflerden keşfetmek, resmetmek ve genişletmek için fırsatlar sunmaktadır [23].

Olay verisi model, durum ve olay olmak üzere üç seviyeden oluşmaktadır [36, 37]. Bu seviyeler Şekil 1'de gösterilmiştir. Her durum ya da olay bir süreç modeline aittir. Durum seviyesinde duruma ait eşsiz durum numarası, durumun demografik özellikleri gibi zamana ve o etkinliğe bağlı değişmeyen özelliklerini içerir. Olay seviyesinde ise eşsiz olay numarası ve durumun süreç boyunca gerçekleştirmiş olduğu etkinliklerin ismi, zaman damgası, etkinliği gerçekleştiren rol bilgisi, kullanılan kaynak bilgisi ve varsa diğer özellikleri içerir. Süreç madenciliği analizinin ilk adımı olan kontrol perspektifinden süreci keşfetmek için, olay düzeyindeki eşsiz olay numarası, etkinlik adı, etkinlik zaman damgası yeterli olacaktır. Her olayın bir duruma, her durumun da bir sürece ait olduğu unutulmamalıdır. Olay düzeyindeki veriler her olayda değişiklik gösterebileceği için özellikle çok perspektifli süreç madenciliği için kritik rol oynarlar. Olay düzeyindeki verilerin süreç madenciliği analizi ile rol, performans ya da veri perspektifi ile süreci keşfetmek mümkün olur.

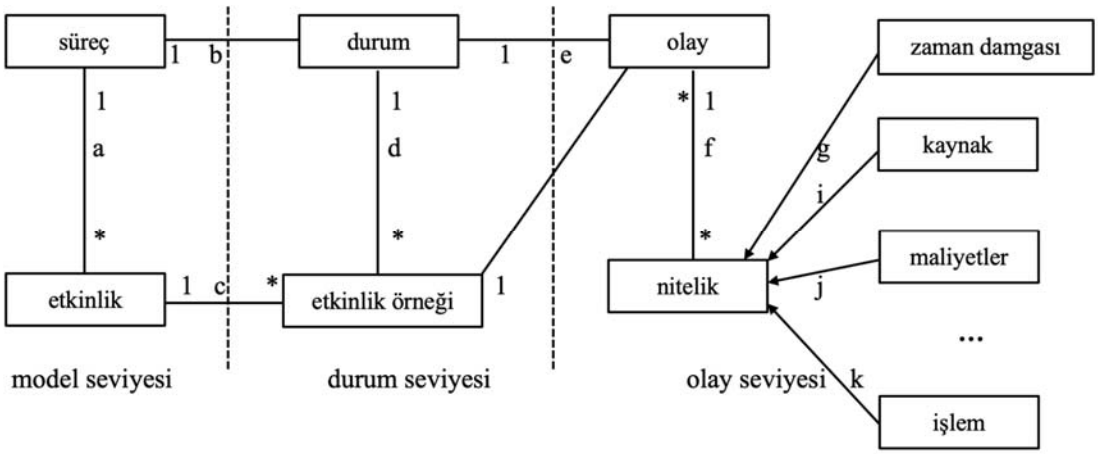
Hastane bilgi sistemleri, hastaya ait tedavi süreçleri, tetkik ve tahlil işlemleri, ödeme bilgileri gibi aslında kendi içinde dağınık ve farklı kaynaklardan beslenen, hemşire, doktor ve teknisyen gibi çeşitli aktörlerin yer aldığı karmaşık sistemlerdir. Hastane yöneticileri ise sağlık süreçlerinin iyileştirilmesi için yalın yönetim ve klinik kılavuzlar gibi süreç iyileştirme yöntemlerini benimseyerek süreçlerini iyileştirme yollarını aramaktadırlar [38, 39]. Sağlık süreç verisi olay günlüğü verisine dönüştürüldüğünde süreç madenciliğinin uygulanabilirliği mümkün olmaktadır.

Sağlık süreçleri verisi genellikle farklı kaynaklardan beslenir. En temel düzeyde incelediğimizde hasta ve hastaya ait demografik bilgiler, muayene gibi hastane ziyaretleri, tetkik ve laboratuvar işlemleri ve müdahale hizmetleri, hastanın geçirmiş olduğu hastalıklar ve ameliyatlar bilgilerini içerir. Muayene tetkik ve müdahale gibi işlemleri gerçekleştiren doktor ya da yetkili bilgisi ile birlikte bu işlemlerin zamanı ve türü sağlık süreçlerinin adımlarını oluşturur. Tedavi süreçleri ve bunların işletimi klinik kılavuzlarda belirtilmiştir. Örneğin, bir acil servis sürecinde hasta muayene edilmeden önce hastanın durumu triyajda değerlendirilmelidir. Bir hastanın on iki saat içinde acilden taburcu edilmesi beklenmektedir. Ameliyat sürecinde servis ya da bölüm değişiklikleri olabilir. Hastaya ait tüm etkinlikler, etkinliklerin gerçekleşme sırası ve etkinlikler arasında geçen süre, sürecin performansını değerlendirmek adına çok önemlidir. Ancak her bir etkinlik, olay seviyesinde hasta ile ilişkilendirildiğinde dağıtık yapıdaki işlemleri birleştirebilmek mümkün olacaktır.

2.2. İlişkili Çalışmalar (Related Work)

Son yıllarda, sağlık sektöründe çeşitli kullanım durumları için süreç madenciliği kullanılmıştır [3, 6]. Bunlar; süreç keşfi, uygunluk kontrolü, süreç iyileştirme, süreç varyasyon analizi, tahmine dayalı izleme ve performans analizidir. Süreç keşfinden sonra en sık kullanım durumunun süreç varyasyon analizinin olması sağlık süreçlerinin çok varyasyonlu diğer bir deyişle dağıtık ve karmaşık yapısının bir göstergesidir. Sağlık hizmetleri süreçleri, çeşitli süreç madenciliği teknikleri kullanılarak farklı amaçlar için birden fazla perspektiften analiz edilebilir. Uygulamalar; onkoloji, cerrahi, acil servis, nörolojik hastalıklar ve kardiyovasküler hastalıklar olmak üzere hemen hemen tüm sağlık uzmanlık dallarında yapılmıştır. Bunlar arasından onkoloji [40, 41] ve acil servis süreçlerini [42, 20] farklı perspektiflerden analiz eden çalışmalar yer almaktadır.

Ekici vd. [24, 43], bütünsel bir modeli keşfetmek adına kontrol akış, veri ve örgütsel perspektifi bir modelde resmeden bir BPMN veri modeli önermiştir ve bu model ile ayaktan hasta tedavi süreci resmedilmiştir. Tiftik vd. [23], her bir perspektif için farklı süreç madenciliği algoritmalarını uygulamayı mümkün kılan bir çatı ve ProM eklentisi geliştirmiştir. Bu çalışmada geliştirilen eklenti kullanılarak Sepsis tedavi süreci keşfedilmiş ve kontrol, performans ve rol perspektifleri bütüncül bir BPMN modeli ile resmedilmiştir.



Şekil 1. Olay verisi modeli [36] (Event data model [36])

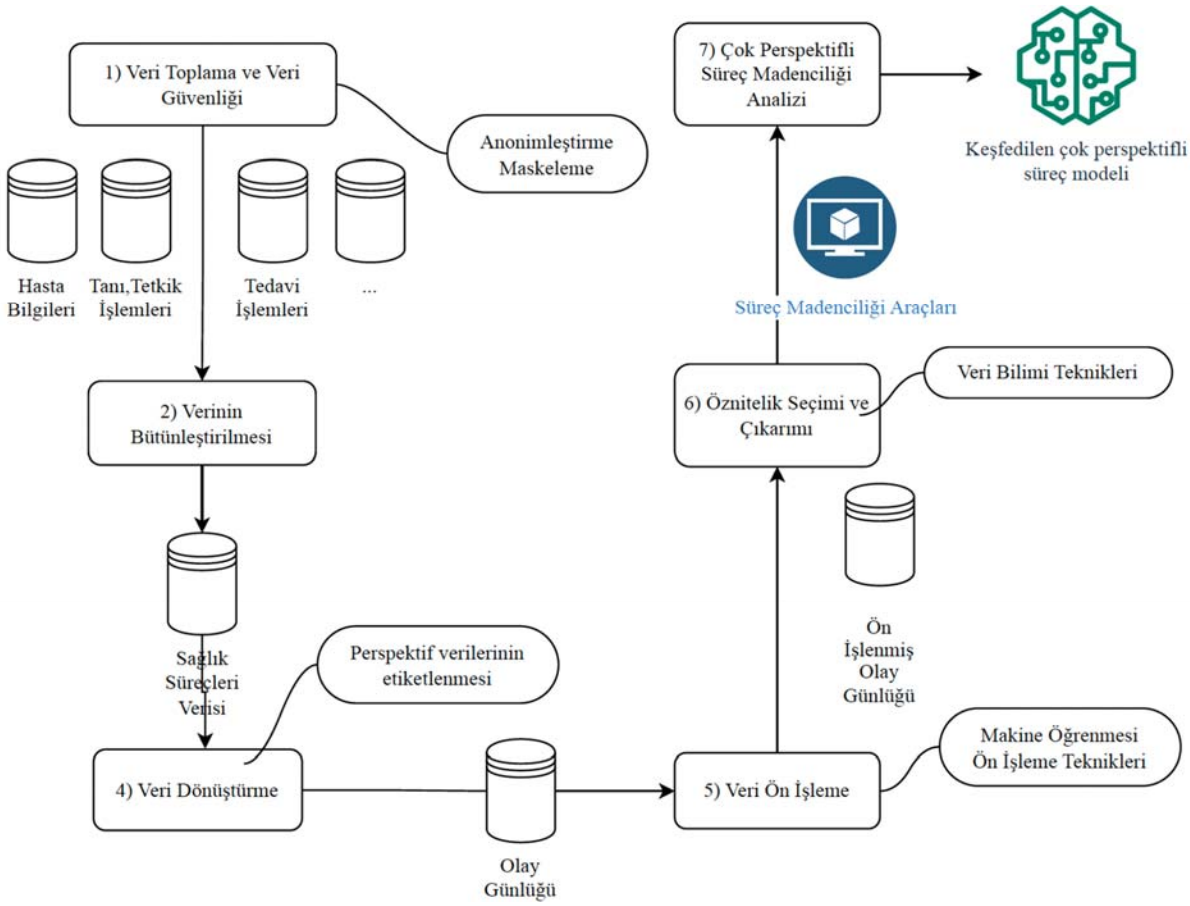
Çalışmaların bir kısmı ise sağlık süreçleri için olay günlükleri oluşturarak, olay günlüklerini onararak, gürültülü verileri ortadan kaldırarak, veri çıkarma, veri modelleme, veri entegrasyonu ve veri ön işleme yoluyla sağlık verilerini değerlendirmenin yollarını keşfetmekle ilgilidir [6]. Sağlık alanında önerilen veri dönüştürme yöntemleri, veri güvenliği konuları üzerinde yoğunlaşmıştır [44]. Önerilen araç ve metodolojiler arasında ise hedef ve soru tabanlı süreç madenciliği projeleri benimsenmiştir [20, 21, 45]. Bu çalışmada önerilen veri dönüştürme yöntemi adımları geliştirilirken, hedef tabanlı süreç madenciliği projesi adımları izlenmiş ve veri güvenliği konuları dâhil edilmiştir. Yöntem, veri bilimi metodolojisindeki veri ön işleme ve özellik seçimi/çıkarımı adımlarında yer alan teknikler ile de zenginleştirilmiştir. Karmaşık ve dağıtık yapıdaki sağlık verisinin bütünleştirilmesi için yollar gösterildiği için de bir kılavuz niteliğindedir. Literatürde sağlık süreçleri için çok perspektifli süreç madenciliği projesinin adımlarını uçtan uca açıklayan ve nihai olarak bütüncül bir süreç modeli sunan ilk çalışmadır. Önerilen yöntemin tüm özellikleri değerlendirildiğinde bütün olarak hem uygulayıcılar için hem de araştırmacılar için çok perspektifli süreç madenciliğinin sağlık alanında uygulamalarını destekleyecek niteliktedir.

3. Yöntem (Method)

Bu çalışmada önerilen veri dönüştürme yönteminin amacı çok perspektifli süreç madenciliği tekniklerini sağlık süreçlerine uygulamak adına, sağlık süreçleri verisini olay günlüğü verisine dönüştürmek için gereken iş adımlarını açıklamaktır. Bu yöntem Şekil 2'de gösterildiği üzere veri toplama ve veri güvenliği, verinin bütünleştirilmesi, veri dönüştürme, veri ön işleme, özellik seçimi ve çıkarımı ve son olarak çok perspektifli süreç madenciliği analizi olmak üzere altı adımdan oluşmaktadır.

3.1. Veri Toplama ve Veri Güvenliği
(Data Collection and Data Privacy)

Hastane bilgi sistemlerinde kaydedilen verinin büyüklüğü her geçen gün daha da artmaktadır. Bir süreç madenciliği projesinde veri toplama, hedefe dayalı yürütülmelidir [19, 21]. Projenin hedefleri, neyi anlamayı, ne için, kim için ve hangi sağlık süreci olduğunu açıkça belirtmelidir. Örneğin acil servis sürecinin performansının hastane yönetici bakış açısı ile anlamaya yönelik bir hedef tanımlanabilir. Bu hedef doğrultusunda veri toplama gerçekleştirilmelidir. Hastane bilgi sistemlerinde yer alan ve hali hazırdaki raporlar, hedef doğrultusunda hazırlanan sorgular ve tıbbi cihazların raporları ile veri toplanmaktadır. Dağıtık yapıdaki her veri için mutlaka eşsiz hasta numarası ve etkinlik numarası, etkinliklere ait saat ve tarih düzeyinde zaman bilgisi toplanmalıdır. Veri toplama esnasında hasta, doktor ve diğer personellere ait bilgilerin toplanması ve saklanması Kişisel Verileri Koruma Yönetmeliğine [46] uygun olarak yapılmalı ve veri güvenliği konuları gözetilmelidir. Süreç madenciliğinde toplanan veri, genellikle bilgi sistemlerinde kaydedilen arşiv verisidir ve etik kurul onayı ile elde edilebilmektedir [47]. Veri gizliliği adına T.C. numarası, ad, soy ad gibi bilgiler veri kümesine dâhil edilmemeli, hastaya, doktora ve diğer personellere ait numaralar ise anonimleştirilmelidir. Yaş gibi diğer kişisel bilgiler projenin hedeflerine bağlı olarak ise genelleştirilebilir. Çok perspektifli süreç madenciliği projesinde ise hedeflenen perspektiflere göre toplanacak verilerin ilgili perspektifle olan ilişkisi açıklanmalı ve gerekirse eşlenmelidir.



Şekil 2. Veri Dönüştürme Yöntemi (Data Transformation Method)

3.2. Verinin Bütünleştirilmesi (Data Integration)

Verinin bütünleştirilmesi, birçok veri kaynağını birleştirilmesidir [48]. Bu adımın sonunda birleştirilmiş bir görünüm sunulur. Sağlık süreçleri verileri, hastane bilgi sistemleri, klinik veri ambarları, gerçek zamanlı veri kaynakları, tıbbi cihazlar (X-ray makineleri), video incelemeleri ve finansal amaçlar için oluşturulan veri kümeleri dahil olmak üzere farklı veri kaynaklarından farklı ayrıntı düzeylerinde toplanmaktadır. Bu kaynaklardan farklı özelliklere sahip sağlık süreçleri verilerini belirlemek ve birleştirmek zorlu görevlerdir.

Farklı kaynaklarda diğer bir deyişle tablo ya da ekranlarda yer alan sağlık süreç verisinin bütünleştirilmesi durum/örnek ve olay seviyesinde olmak üzere iki boyutta gerçekleştirilmektedir. Sağlık sürecinde durum nesnesi hastadır. Durum seviyesinde hastaya ait eşsiz numara üzerinden hastanın durum seviyesinde tüm özellikler tek bir listede toplanmalıdır. Bu değerler olay düzeyinde de aynı değere sahip olabilirler ve olay düzeyindeki veri yapısında da yer alacaklardır. Durum düzeyindeki özellikler: hasta yaşı, cinsiyet, uyruk bilgisi, hastaneye geliş şekli, tanı, hastanın komorbidite bilgileri, başvurulan bölüm ve servis vb. olabilir [19]. Durum düzeyindeki veri özellikleri olay düzeyinde veri özellikleri olarak tekrar edebilir. Toplanan sağlık süreci verisinde olay düzeyindeki veriler açıkça yer almayabilir. Bu düzeyde oluşturulacak etkinlik bilgileri, süreç madencisi ve alan uzmanının birlikte çalışması ile elde edilmelidir.

Hastanın hastaneye yatışı süresi, taburcu süresi gibi zaman damgaları toplanan veride yer alıyorsa, olay düzeyinde olay günlüklerini oluşturmak için bu zaman damgalarıyla eşleşecek etkinlik ismi süreç madencisi tarafından alan uzmanı ile birlikte belirlenmelidir. Büyük bir çoğunlukla toplanan sağlık verisinde yer almayan etkinlik isimleri, saat ve tarih düzeyinde zaman damgası bilgisi ve klinik kılavuzlar incelenerek oluşturulmalıdır. Oluşturulan etkinlik isimleri, etkinlik zamanları ve bu etkinlikleri gerçekleştiren hastane personeli bilgisi ve diğer ilişkili veri özellikleri olay günlüklerini oluşturur. Burada dikkat edilmesi gereken önemli hususlardan biri, her bir olayın eşsiz olay numarası ve her bir etkinliğin eşsiz etkinlik numarası veri bütünleştirme adımında oluşturulmalıdır.

Özetle veri bütünleştirme adımında süreç madencisinin dikkat etmesi gereken hususlar; farklı kaynaklardaki verileri birleştirmek için eşsiz hasta numarası izlenmeli, olay düzeyindeki etkinlik isimleri etkinlik zaman bilgileri ve alan bilgisi değerlendirilerek oluşturulmalı, eşsiz olay numarası ve etkinlik numarası genel sabit bir dizi değişkeni kullanılarak oluşturulmalıdır.

3.3. Veri Dönüştürme (Data Transformation)

Veri dönüştürme, birleştirilen verinin çok perspektifli süreç madenciliği tekniklerini uygulamaya hazır hale getirilmesi için yapılan bir dizi işlemlerden oluşur [34]. Hasta tanımlayıcı özellik olan hasta numarası üzerinden belirlenen durum ve olay düzeyindeki veri özellikleri, olay düzeyinde olay günlüğüne dönüştürülmektedir.

Dönüştürme adımı için olay veri yapısı (Şekil 1) temel alınır. Dönüştürme işlemi iki adımdan oluşur. İlk adımda hastaya ait durum düzeyindeki veriler belirlenir. Bu veriler rol, performans ve veri perspektifi özellikleri olarak etiketlenir. Örneğin, hastanın yaşı, cinsiyeti, uyruğu gibi demografik özellikleri veri perspektifi verileri olarak etiketlenir. Başvurduğu bölüm, muayene ya da ameliyat doktoru eğer bu özellikler ilgili sağlık süreci boyunca aynı değerde olacaksa kaynak ya da rol perspektifi verileri olarak etiketlenir. İkinci adımda ise toplanan veride yer alan birbirinden farklı zaman damgaları o hastaya ait olay verilerini oluşturur. Tüm zaman damgaları için farklı etkinlik adı belirlenir. Olay düzeyindeki veri, durum düzeyindeki veri özelliklerini içerecek şekilde ve olaya ait eşsiz olay numarası, zaman damgası, etkinlik adı ve diğer olay düzeyinde veriler kullanılarak oluşturulur.

Bu aşama sonunda bütünleştirilen ve hasta tabanlı sağlık verisi olay günlüğü verisine dönüştürülmüş olur. Oluşturulan olay günlüğü verisinin formatının doğruluğu; ProM [27], Disco [28], PM4Py [29] ya da RBupar [31] süreç madenciliği araçları kullanılarak test edilmelidir.

3.4. Veri Ön İşleme (Data Preprocessing)

Bu aşamada eksik, hatalı, gürültülü değerler ve tekrarlar içeren düşük kaliteli olay günlükleri ön işleme adımlarından geçirilir. Ön işleme adımları, süreç madenciliği analizinin performansını arttırmak ve daha anlaşılır bir model oluşturmak amacıyla uygulanır. Olay verileri ön işleme adımları herhangi bir makine öğrenmesi projesinde yapılan veri ön işleme adımlarıyla benzerdir [35, 49, 50]. Veri temizleme işlemleri ile birlikte iz kümeleme ve iz/olay düzeyinde filtreleme en sık kullanılan ön işleme teknikleri arasındadır [51]. Veri temizleme adımları, boş verilerin tespiti, bu verilerin çıkarılması ya da ortalama değerler ile doldurulması, genellikle metin temizleme adımlarıdır. İz kümelemenin ve iz/olay düzeyinde filtreleme teknikleri ise çoğu süreç madenciliği aracı ile desteklenmesi, uygulama bakımından kolay olması ve olay günlüklerindeki gürültüyü ve eksikliği yeterince yönettiği için tercih sebebi olmuştur.

3.5. Öznitelik Seçimi ve Çıkarımı (Feature Selection and Extraction)

Öznitelik seçimi, fazla bilgi kaybına uğramadan ilgili özelliklerin ilgili alt kümelerini seçmeye çalışır [52] ve öznitelik çıkarma teknikleri, daha az yedekli yeni bir değişken kümesi elde etmek için orijinal öznitelik kümesini birleştirir [53]. Öznitelik seçimi ve çıkarımı veri bilimi metodolojisinde üretilen modelin kalitesini doğrudan etkileyen tekniklerdir.

Ön işleme adımları ile elde edilen yüksek kaliteli olay günlükleri, daha yalın bir veri tipine dönüştürülmektedir. Böylece daha anlaşılır bir model keşfetme mümkün olacaktır. Ön işleme sonrasında durum ve olay seviyesindeki öz nitelikler hem alan bilgisi göz önünde bulundurularak hem de özellik seçimi teknikleri kullanılarak *filtreleme* ile belirlenmektedir. Filtreleme teknikleri; entropi tabanlı filtrelemeyi, doğrusal özellikleri filtrelemeyi ve frekans tabanlı filtrelemeyi içerir [54]. Diğer öznitelik seçimi teknikleri ise *gömülü* ve *sarma* olmak üzere iki ana yaklaşıma ayrılmaktadır [55]. Öznitelik seçimi, tahmine dayalı izleme ve veri perspektifinden sürecin keşfedilmesi için kritik rol oynar.

Öznitelik çıkarımı, olay günlüğünün çok perspektifli süreç madenciliği ile farklı perspektiflerden keşfedilmesi sonucu elde edilen çıktıların kullanımı ile olmaktadır. Döngüler halinde oluşturulan yeni nitelikler, bir başka perspektifin keşfedilmesi veya modelin daha anlaşılır olması için girdi olmaktadır. Örneğin, kontrol perspektifi ile sürecin keşfedilmesi sonucu oluşan iz dizilimleri ya da hastanın süreç boyunca hastanede bulunduğu toplam süre bilgisi hesaplanır ve yeni bir öznitelik olarak elde edilir. Bu işleme olay günlüğünün zenginleştirilmesi adı verilmektedir [32]. Zenginleştirilen olay günlüğünün tekrar süreç madenciliği teknikleri ile analiz edilmesi; süreçteki problemlerin, istisna durumların ve darboğazların nedenlerinin araştırılmasını desteklemektedir.

3.6. Çok Perspektifli Süreç Madenciliği Analizi (Multi-perspective Process Mining)

Sağlık verisinin çok disiplinli ve karmaşık yapısı dolayısıyla sürecin sadece kontrol perspektifi ile keşfedilmesi yeterli olmamaktadır [22]. Literatürde de sağlık süreçlerine en sık uygulanan süreç madenciliği

teknikleri arasında süreç keşfi, süreç, uygunluk kontrolü, süreç iyileştirmeye ek olarak varyasyon analizi, performans analizi gibi teknikler yer almaktadır [6]. Bu bulgular süreci farklı perspektifler ile keşfetmeyi diğer deyişle keşfedilen süreci performans, rol, veri gibi perspektiflerle geliştirme gerekliliği destekler. Bu nedenle önerilen yöntemdeki süreç madenciliği analizi teknikleri süreç keşfi, süreç varyasyon analizi, performans analizi teknikleri olarak belirlenmiştir.

4. Durum Çalışması (Case Study)

Sağlık verisinden olay verisine dönüştürme işlemlerinin açıklandığı yöntem, Türkiye'de yer alan bir üniversite hastanesine ait ameliyat süreci verisi üzerinden durum çalışması ile doğrulanmıştır. Durum çalışması protokolü olarak yazılım mühendisliği alanında yaygın olarak kullanılan durum çalışması tasarım yaklaşımları uygulanmıştır [26].

Tek ve gömülü durum çalışması, keşifsel çalışma olarak tasarlanmıştır. Durum çalışması tek bir hastanenin verisini kullanır ve ameliyat sürecinde farklı gruplardaki ameliyatlar yer aldığı için çalışma gömülüdür. Önerilen yöntemi doğrulamak amacıyla yapıldığı için keşifsel durum çalışmasıdır. Durum çalışmasının yapıldığı üniversite hastanesi klinik yol ve kılavuzları ve kanıta dayalı tıp yaklaşımını benimsemiştir. Hastane bilgi sisteminde 2017 yılı boyunca kaydedilen tüm ameliyat süreci verisini kapsar ve büyük veri olarak nitelendirilebilir. Veri kaynağı üçüncü derece veri tipi olan arşiv verisidir. Bu veriler çalışma için kaydedilmemiştir ama veri çıkarımı hedefe yönelik yapılmıştır. Çalışmanın hedefi hastanenin ameliyat sürecinin zaman verimliliğini ölçmektir. Veri çıkarımı veri tabanı sorgusu yazılarak gerçekleştirilmiştir.

Veri dönüştürme yöntemi uygulanarak sürece ait temel istatistiki bilgiler de edinilmiştir. Diğer bir deyişle önerilerin yöntem bir veri keşfetme yöntemidir. Durum çalışması ise önerilen yöntemin adımları izlenerek yürütülmüştür. Veri dönüştürme işlemleri için PM4PY ve RBupar, çok perspektifli süreç madenciliği analizi için ise ProM ve Disco araçları kullanılmıştır.

- *Veri Toplama ve Veri Güvenliği*

Ameliyat sürecinin zaman verimliliğini ölçmek adına 2017 yılı boyunca tüm ameliyat grupları için toplanan veri, hastane bilgi işlem görevlisi tarafından oluşturulan sorgu ile hastane bilgi sisteminden otomatik olarak çekilmiştir. Veri, hastanın demografik bilgileri, geçirdiği ameliyatlar ve ameliyatlar ile ilgili özellikler ve ameliyat süreci ile ilgili zaman bilgilerini içermektedir. Tüm zaman bilgileri ameliyat istek tarihi, ameliyat kabul tarihi, varsa ameliyat iptal tarihi, yatış tarihi, varsa servis değişiklik tarihi, bölüm değişiklik tarihi, ameliyat başlangıç tarihi, bitiş tarihi ve taburcu tarihidir. Zaman verileri tarih ve saat olarak yer almaktadır. Süreç madenciliği için zaman damgaları, sürecin uçtan uca resmedilmesi için kritik özelliklerdir [1]. Ayrıca hastaya ameliyat öncesinde yapılan muayene zamanı ve tetkik bilgileri ile hastayı muayene eden doktor tarafından konulan teşhis bilgisi de yer almaktadır. Ameliyat süreci verisi; hasta bilgileri, doktor bilgileri, ameliyat bilgileri ve muayene bilgileri olmak üzere dört temel kaynaktan elde edilmiştir.

- *Verinin Bütünleştirilmesi*

Anonimleştirilmiş ve dört farklı kaynakta yer alan ameliyat süreci verisi hasta numarası üzerinden tek bir liste olarak birleştirilmiştir. Bütünleştirilen verinin birincil anahtarı, tablodaki her satırı benzersiz şekilde tanımlamak adına hasta numarası alanıdır. Hastane bilgi sistemlerinde hasta numarası, eşsiz değerler içerdiği, boş ve sıfır değer içermediği ve içerdiği değerler hiç değişmediği için birincil anahtar olmaya uygundur.

Bir hasta birden fazla ameliyat geçirebildiği için bir hastaya ait birden fazla satırın yer alması olağandır. Ama tekrar eden verileri bu aşamada elemek önemlidir. Bütünleştirme esnasında alan uzmanı ile birlikte çalışmanın hedefine yönelik özellik seçimi de yapılabilir.

Bu durum çalışmasında bütünleştirilen veri, tekrar edilen verilerin de çıkartılması ile 11975 hastaya ait 19064 satır ve 29 kolondan oluşur ve veri seti ameliyat süreci sağlık verisi olarak adlandırılmıştır. Dört hastaya ait ameliyat süreç verisinin bir kısmı Tablo 1'de gösterilmiştir. Verinin alanları hasta id, hasta no, hastanın yaşı, yatış tarihi, taburcu tarihi, ameliyatın adı, ameliyat doktoru, ameliyat başlangış ve bitiş tarihi ve ameliyatın grubu (G)'dır. Büyük ve çok boyutlu olan bu sağlık veri setinin, veri ön işleme ve özellik seçimi teknikleri ile boyutunun indirgenme ihtiyacı kaçınılmazdır. Bu teknikler daha basit ve daha anlaşılır süreç modelleri oluşturmayı, süreç madenciliği tekniklerinin performansını iyileştirmeyi ve temiz, anlaşılır veriler hazırlamayı sağlayacaktır [1].

- *Veri Dönüştürme*

Veri dönüştürme aşamasında RbupaR [31] aracındaki eventLog kütüphanesi kullanılmıştır. İlk adım olarak olay ve etkinlik örneği numarası için sabit bir değişken belirlenmiştir. Ameliyat sağlık süreci verisinin her satırında, her zaman damgası için bir olay verisi oluşturulmuştur. Zaman damgalarının sıralaması sürecin adımları oluşturacaktır. Veri dönüştürme işleminin, süreç madencisinin alandaki uzmanlığına bağlı olarak bir alan uzmanı ile birlikte

Tablo 1. Bütünleştirilmiş ve anonimleştirilmiş ameliyat süreci sağlık verisi
(Integrated and anonymized surgery process healthcare data)

Hasta ID	Hasta No	Yaş	Yatış Tarihi	Taburcu Tarihi	Ameliyat Adı	Doktor	Ameliyat Baş. Tarihi	Ameliyat Bitiş Tarihi	G
68	111466	76	01/02/2017 13:57:00	02/02/2017 17:37:00	Biyopsi, deri veya derialtı, yüzeysel	R1	01/02/2017 16:20:00	01/02/2017 16:50:00	E
164	122473	72	15/09/2017 07:48:00	18/09/2017 07:35:00	FAKO + İOL	R2	14/08/2017 14:34:00	14/08/2017 14:00:00	B
344	126904	64	06/11/2017 08:06:00	06/11/2017 09:47:00	Kafatası kemiği/saçlı derinin basit tümöral kitleleri	R3	06/11/2017 08:30:00	06/11/2017 09:30:00	C
366	121923	72	31/07/2017 10:10:00	07/08/2017 10:45:00	Büyük eklem parsiyel protezleri, primer	R4	03/08/2017 13:30:00	03/08/2017 16:00:00	B
669	112573	58	05/02/2017 15:07:00	21/02/2017 10:44:00	Artroskopik Akromioplasti, omuz	R5	20/02/2017 12:00:00	20/02/2017 15:00:00	B
849	125655	72	13/10/2017 08:16:00	14/10/2017 11:10:00	Üreteral J Stent takılması, değiştirilmesi	R6	13/10/2017 10:40:00	13/10/2017 12:00:00	C

yapılması da gerekebilir. Alan uzmanı ile birlikte çalışma imkânı olmazsa sürece ait klinik yol ve klinik kılavuzlar ve toplanan anahtar performans göstergeleri de sürecine adımlarını oluşturmak adına faydalı olacaktır [6, 35].

Süreç madenciliği tekniklerini alan bilgisi olmadan gerçekleştirmek mümkün değildir. Bu durum çalışmasında da hastanenin ameliyat süreci klinik yol ve kılavuzları incelenmiş, ölçülen anahtar performans göstergeleri ve değerleri de göz önünde bulundurularak etkinlik isimleri oluşturulmuştur.

Oluşturulan olay günlüğü verisi 68 numaralı hasta için Tablo 2'de listelenmiştir. 68 numaralı hastanın ameliyat istek kaydının girilmesi ile ameliyat süreci başlamıştır. Ameliyat istek talebi iptal edilebilir ya da hasta hastaneye yatış yapar. Yatış sonrasında ameliyat başlar ve ameliyat bittikten sonra servise kaldırılır. Ameliyat bitiş kaydı olmadan hasta servise kaldırılmamalıdır. Servise kaldırılan hasta ise taburcu olmalıdır. Hastaneye yatış ve taburcu esnasında servis veya bölüm değişikliği olabilir. Yine hastanın ameliyat süreci boyunca hasta vefat edebilir. 68 numaralı hasta 01/02/2017 13:57 tarihinde hastaneye yatmış ve 02/02/2017 17:37 tarihinde taburcu olmuştur. Ameliyat sürecinin yönetimi adına önemli bir gösterge olan yatış süresi 68 numaralı hasta için yaklaşık olarak 1 gün 4 saattir.

- *Veri Ön İşleme & Öznitelik Seçimi ve Çıkarımı*

29 farklı özelliği olan sağlık süreci verisinde yer alan özelliklerin boş değer içeren satır sayıları ile hesaplanmıştır. Boş değer içeren tarih alanları süreç madenciliği analizine dâhil edilmemiştir. Bu alanlardaki boş veri ayrıca alanla ilgili etkinliğinde gerçekleşmediğinin göstergesi olmuştur. Örneğin, ameliyat istek iptal tarihi alanı boş ise ameliyat talebinin iptal olmadığını ve bu etkinlikle ilgili bir olay kaydının oluşturulmayacağını gösterir. Ayrıca sabit değer içeren alanlarda çıkarılmıştır. Sağlık süreci verisi için elde edilen olay günlüğü verisi 121847 satır ve 14 kolondan oluşmaktadır ve bir kısmı Tablo 2' de gösterilmiştir. Alınan bu veriler yansız verilerdir, açıklamalar doğrultusunda önerilen yöntemin adımları izlenerek veri dönüştürme aynı sonuçlar ile tekrarlanabilir. Ayrıca keşfedilen sürecin geçerlilik ve güvenilirliği ProM, Disco ve RBupar olmak üzere üç farklı süreç madenciliği aracı ile test edilmiştir. Disco aracının süreç modeli daha anlaşılır olduğu tercih edilmiştir.

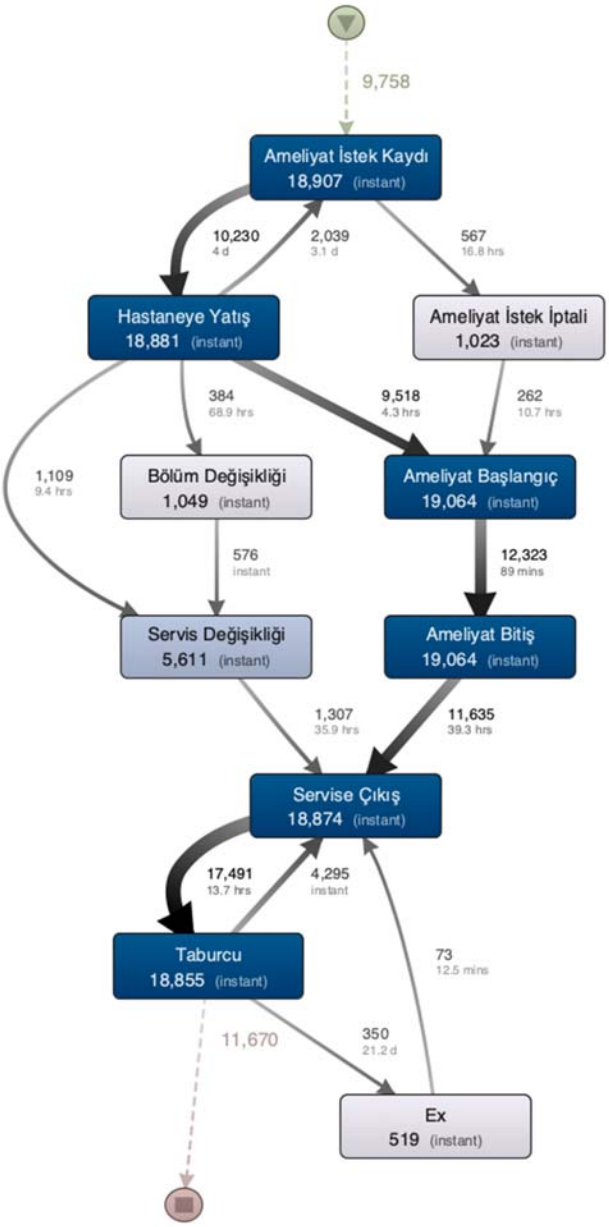
- *Çok Perspektifli Süreç Madenciliği Analizi*

Ameliyat süreci olay günlüğü verisi Disco süreç madenciliği aracı kullanılarak kontrol ve performans perspektifleri birleştirilerek keşfedilmiştir. Keşfedilen süreç Şekil 3'de yer almaktadır. Kontrol perspektifini incelediğimizde, koyu mavi ile gösterilen etkinlikleri ve sırası süreçte izlenen en sık yoldur. Ameliyat süreci toplamdan 11.975 hastanın yer aldığı ve 18.907 hastanın ameliyatının planlanmasıyla "Ameliyat İstek Kaydı" ile başlar, "Hastaneye Yatış", "Ameliyatın Başlangıç", "Ameliyat Bitiş", "Servise Çıkış" ve "Taburcu" ile sonlanır. Taburcu edilen hasta sayısı 18.855'tir. Bu etkinlikler dizilimi zaten ameliyat süreci için olağan senaryodur. Bunun yanında ameliyat planlanmasında bir sorun olduğunda "Ameliyat İstek İptali" etkinliği gerçekleşebilir. Yine kaynak düzenlenmesi ya da başka bir nedenle "Servis Değişikliği" ya da "Bölüm Değişikliği" etkinlikleri gerçekleşebilir.

Bu senaryolar alternatif senaryo olarak adlandırabilir ve sonuçlarında maliyet ve zaman açısından kayıplar yaşanabilir. Alternatif senaryolarının nedenlerini araştırmak süreç iyileştirme fırsatlarını doğuracaktır. Performans perspektifi için ise her etkinlik arasında geçen ortalama zaman süreleri yine Şekil 3'te gösterilmiştir. Örneğin, hastanın yatışından ortalama 4 gün öncesinde ameliyat planlanmasının yapıldığı gözlemlenmektedir. Bu ise hastane yönetimi açısından kritik ve başarılı bir gösterge değeri olarak söylenebilir. Ortalama ameliyat süreci 89 dakika olarak ölçülmüştür. Bu değeri ameliyat gruplarına göre hesaplamak daha doğru olacaktır. Ayrıca ameliyatın bitişinden servise çıkış arasında geçen ortalama süre değeri 33 saat olarak ölçülmüştür. Bu değer beklenenden uzun bir bekleme süresidir. Sağlık süreçlerinin doğası gereği bazı verilerin uygulamayı yansıtmadığı ve işlem yapıldıktan sonra veri girişi yapıldığı tahmin edilse de böyle bir gecikmenin nedenleri araştırılmalıdır.

Tablo 2. Ameliyat süreci sağlık verisine ait olay günlüğü verisi (Event logs for surgery process data)

Hasta ID	Olay ID	Ameliyat No	Yaş	Etkinlik	Tarih	Ameliyat Adı	Ameliyat Doktoru	Bölüm Adı	Ameliyat Grubu
68	1	111466	76	Ameliyat İstek Kaydı	31/01/2017 14:17:36	Biyopsi, deri veya derialtı, yüzeysel	R1	KBB	E Grubu
68	2	111466	76	Hastaneye Yatış	01/02/2017 13:57:00	Biyopsi, deri veya derialtı, yüzeysel	R1	KBB	E Grubu
68	3	111466	76	Ameliyat Başlangıç	01/02/2017 16:20:00	Biyopsi, deri veya derialtı, yüzeysel	R1	KBB	E Grubu
68	4	111466	76	Ameliyat Bitiş	01/02/2017 16:50:00	Biyopsi, deri veya derialtı, yüzeysel	R1	KBB	E Grubu
68	5	111466	76	Taburcu	02/02/2017 17:37:00	Biyopsi, deri veya derialtı, yüzeysel	R1	KBB	E Grubu
164	6	122473	72	Ameliyat İstek Kaydı	10/08/2017 10:50:08	FAKO + İOL	R2	Göz Hastalıkları	B Grubu
164	7	122473	72	Hastaneye Yatış	14/08/2017 11:59:00	FAKO + İOL	R2	Göz Hastalıkları	B Grubu
164	8	122473	72	Ameliyat Başlangıç	14/08/2017 14:00:00	FAKO + İOL	R2	Göz Hastalıkları	B Grubu
164	9	122473	72	Ameliyat Bitiş	14/08/2017 14:20:00	FAKO + İOL	R2	Göz Hastalıkları	B Grubu
164	10	122473	72	Servise Çıkış	14/08/2017 14:55:00	FAKO + İOL	R2	Göz Hastalıkları	B Grubu
164	10	122473	72	Taburcu	14/08/2017 14:55:00	FAKO + İOL	R2	Göz Hastalıkları	B Grubu



Şekil 3. Keşfedilen ameliyat süreci modeli
(Discovered surgery process model)

5. Sonuçlar (Conclusions)

Bir iş süreci yönetim tekniği olan süreç madenciliğinin sağlık alanında uygulamaları her geçen gün artmaktadır. Süreç madenciliğinin sağlık alanında uygulamaları arasında en sık uygulanan teknikleri süreç keşfi, uygunluk kontrolü, süreç varyasyon analizi, performans analizi ve tahmine dayalı izlemedir. Bu çalışmada süreç madenciliğinin ilk adımı olan süreç keşfi için sağlık verisinin olay verisine dönüştürme adımlarını açıklayan bir yöntem önerilmiştir. Önerilen veri dönüştürme yöntemi veri toplama ve veri güvenliği, verinin bütünleştirilmesi, veri dönüştürme, özellik seçimi ve çıkarımı ve süreç madenciliği olmak üzere altı adımdan oluşmaktadır. Önerilen yöntem sıradan süreç madenciliği veri dönüştürme yöntemlerinden farklı olarak sağlık verisi özelinde geliştirilmiş ve veri güvenliği konuları, veri bilimi metodolojisi ve veri bilimi teknikleri ile zenginleştirilmiştir. Yöntem sağlık verisinin doğası gereği içerdiği karmaşıklık, dağıtık ve çok disiplinli yapısındaki zorlukları ele almak için pratik bir örnek içererek yol gösterici yaklaşımlar içermektedir. Hedef tabanlı süreç madenciliği projesinin taban alınması, süreç madenciliği araçlarından örnekler verilmesi ve olay günlüğünün olay-veri modelinin seviyeleri düzeyinde oluşturulması bu yaklaşımlar arasındadır. Olay günlüklerinin oluşturulmasından dönüştürülmesine kadar tüm aşamaların tek bir çalışmada yer alması, çok perspektifli süreç madenciliği için verilerin perspektiflerle ilişkilendirilmesi ve bu perspektifler ile resmedilmesi çalışmanın özgün katkılarındandır.

Veri dönüştürme yöntemi Türkiye'deki bir üniversitesi hastanesine ait ameliyat süreci verisi ile doğrulanmıştır. Yapılan durum çalışmasında ameliyat sürecine ait sağlık verisi toplanmış, veri güvenliği konuları ele alınarak veri anonimleştirilmiştir. Anonimleştirilen ve dört farklı kaynaktan ele edilen sağlık verisi, hasta odaklı olarak birleştirilmiştir. Hasta odaklı veri olay günlüğü verisine dönüştürülmüştür. Veri dönüştürme esnasında alan bilgisi de kullanılarak etkinlik isimleri belirlenmiştir. Dönüştürülen verinin çok boyutlu olması nedeniyle 29 özellik arasından 14 özellik seçilmiş ve verinin karmaşıklığı azaltılmıştır. Bu özellikler alan uzmanı ile birlikte hastanenin süreç tanımı, ölçülen anahtar performans göstergeleri baz alınarak seçilmiştir.

Sonuç olarak ise oluşturulan olay günlüğü verisine Disco süreç madenciliği aracı kullanılarak süreç keşfi algoritması uygulanmıştır. Keşfedilen süreç modeli iki farklı perspektif birleştirilerek resmedilmiştir. Nesne tabanlı ameliyat süreci verisinin, olay günlüğüne dönüştürülmesi için gereken tüm adımlar detaylıca açıklanmış ve bu işlemler pratik bir örnek ile gösterilmiştir. Farklı kaynaklardan farklı özelliklere sahip sağlık süreçleri verileri genellikle eksiktir ve süreç merkezli olmaktan çok nesne merkezlidir ve farklı ayrıntı düzeylerinde aykırı değerler ve olaylar içerir [6]. Çalışma ve önerilen yöntem, sağlık hizmetleri alanında süreç madenciliği tekniklerinin etkin kullanımı için veri modelleme, çıkarma, bütünleştirme, ön işleme, görselleştirme ve inceleme ile ilgili sorunların ele alınmasına ve çözülmesine katkıda bulunmaktadır.

Gelecekteki çalışmalarda, önerilen yöntemi daha fazla durum çalışması yaparak diğer sağlık süreçlerinden acil süreçlerine uygulanması planlanmaktadır. Ayrıca, önerilen yöntemin değerlendirme yöntemini ve operasyonel senaryoların yürütülmesini ve çok perspektifli süreç madenciliği özelliklerini destekleyecek bir aracın geliştirilmesine ihtiyaç vardır. Süreç keşfi, süreç karşılaştırması, bütünsel bir modelin oluşturulması ve aykırı değer tespiti özelliklerini içerecektir. Böyle bir aracın geliştirilmesi, sağlık hizmetleri süreci verilerinin tıp uzmanları ve hastane yöneticileri tarafından yerinde analizini desteklemek için de faydalı olacaktır.

Kaynaklar (References)

1. Aalst, W. M. P., Process Mining: Data Science in Action, Springer, Heidelberg, 2016.
2. Aalst, W. M. P., Adriansyah A., Medeiros A. K. A. D., Arcieri, F., Baier T., and T. B. et al., Process mining manifesto, Business Process Management (BPM) Workshops: 2011 International Workshops, Clermont-Ferrand-France, 169–194, 2012.
3. Munoz-Gama J. et al., Process mining for healthcare: Characteristics and challenges, J Biomed Inform, 127, 103994, 2022.
4. E. De Roock and N. Martin, Process mining in healthcare – An updated perspective on the state of the art, J Biomed Inform, 127, 103995, 2022.
5. Man's R. S., Aalst W. M. P. Van Der, and Vanwersch R. J. B., Process Mining in Healthcare Evaluating and Exploiting Operational Healthcare Processes, 1-91, Springer International Publishing, Heidelberg, 2015.
6. Erdogan T. and Tarhan A., Systematic Mapping of Process Mining Studies in Healthcare, IEEE Access, 6, 24543-24567, 2018.
7. Mannhardt F., Multi-perspective Process Mining, Doktora tezi, Eindhoven Teknoloji Üniversitesi, Matematik ve Bilgisayar Bilimleri, 2018.

8. Peterson J., Petri Net Theory and the Modeling of Systems, Prentice Hall, NJ, A.B.D., 1981.

9. Weijters A. J. M. M. and Ribeiro J. T. S., Flexible heuristics miner (FHM), IEEE symposium on computational intelligence and data mining (CIDM), Paris-Fransa, 310-317, 2011.

10. Günther CW. and W. Van Der Aalst, Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics, Business Process Management - Lecture Notes in Computer Science, 4714, 328–343, 2007.

11. Object Management Group (OMG), Business Process Model and Notation (BPMN) Version 2.0, www.omg.org/spec/BPMN/2.0.2/, 2013, Erişim Tarihi Aralık 12, 2023.

12. Leoni M. De and Aalst W. M. P. Van Der, Data-aware process mining: Discovering decisions in processes using alignments, Proceedings of the ACM Symposium on Applied Computing, 1454–1461, 2013.

13. Leemans S. J. J., Fahland D., and Aalst W. M. P. Van Der, Process and deviation exploration with inductive visual miner, 12th International Conference on Business Process Management, İsrail, 1295, 46–50, 2014.

14. Doğan O., Overview of Process Mining: Alpha Algorithm for Process Flow Discovery, Pamukkale University Journal of Engineering Sciences, 26 (5), 966–973, 2020.

15. Man's R., Schonenberg M., Song M., Aalst W. Van der, and Bakker P., Application of process mining in healthcare – A case study in a Dutch hospital, International Joint Conference on Biomedical Engineering Systems and Technologies, Portekiz, 425–438, 2008.

16. Orellana Garcia A., Perez Ramirez Y. E., and Armenteros Larrea O. U., Process Mining in Healthcare: Analysis and Modeling of Processes in the Emergency Area, IEEE Latin America Transactions, 13, 5, 1612–1618, 2015.

17. Rovani M., Maggi F. M. Maggi, Leoni M. de, and Aalst W. M. P. van der, Declarative process mining in healthcare, Expert Syst Appl, 42, 23, 9236–9251, 2015.

18. Delias P., Doumpos M., Grigoroudis E., Manolitzas P., and Matsatsinis N., Supporting healthcare management decisions via robust clustering of event logs, Knowl Based Syst, 84, 203–213, 2015.

19. Erdogan T. G. and Tarhan A., A goal-driven evaluation method based on process mining for healthcare processes, Applied Sciences (Switzerland), 8, 6, 2018.

20. Rebuge Á. and Ferreira D., Business process analysis in healthcare environments: A methodology based on process mining, Information Systems, 37, 2012.

21. Eck M. L. Van, Lu X., Leemans S. J. J., and Aalst W. M. P. Van Der, PM 2 : a Process Mining Project Methodology, International conference on advanced information systems engineering, 297-313, Cham: Springer International Publishing, 2015.

22. Erdogan T. G. and Tarhan A. K., Multi-perspective Process Mining for Emergency Process, Health Informatics J, 1–18, 2022.

23. Tiftik M. N., Erdogan T. Gurgen, and Kolukisa Tarhan A., A framework for multi-perspective process mining into a BPMN process model, Mathematical Biosciences and Engineering, 19, 11, 11800–11820, 2022.

24. Ekici B., Erdogan T. G., and KoukisaTarhan A. K., BPMN Data Model for Multi-Perspective Process Mining on Blockchain, International Journal of Software Engineering and Knowledge Engineering, 1–29, 2022.

25. Yin R. K., Case Study Research. Design and Methods, 5, 2009.

26. Runeson P. and Höst M., Guidelines for conducting and reporting case study research in software engineering, Empir Softw Eng, 14 (2), 131–164, 2009.

27. Process Mining Group, ProM - the leading process mining toolkit, 2014, Erişim Tarihi Aralık 12, 2023.

28. Process Mining and Automated Process Discovery Software for Professionals - Fluxicon Disco, Erişim Tarihi Aralık 12, 2023.

29. A. Berti, S. J. Van Zelst, W. M. P. Van Der Aalst, and F. Gesellschaf, Process mining for python (PM4py): Bridging the gap between process-And data science, CEUR Workshop Proc, 2374, 13–16, 2019.

30. About Celonis | The Leader & Innovator in Process Mining. https://www.celonis.com/company/, Erişim Tarihi Aralık 12, 2023.

31. Janssenswillen G., Depaire B., Swennen M., Jans M., and Vanhoof K., bupaR: Enabling reproducible business process analysis, Knowl Based Syst, 163, 927–930, 2019.

32. Suriadi S., Ouyang C., Aalst W. M. P. Van Der, and Hofstede A. H. M., Root cause analysis with enriched process logs, Lecture Notes in Business Information Processing, 132 LNBIP (January), 174–186, 2013.

33. Mannhardt F., Leoni M. De, and Reijers H. A., The multi-perspective process explorer, CEUR Workshop Proc, 1418, 130–134, 2015.

34. Reijers F. Mannhardt and Blinde D., Analyzing the trajectories of patients with sepsis using process mining, CEUR Workshop Proc, 1859, 72–80, 2017.

35. Aalst Wil M. P. van der, Process Mining Handbook, in Lecture Notes in Business Information Processing, 448, Cham: Springer International Publishing, 448, 2022.

36. Erdogan T. G., A Performance Analysis Method for Healthcare Process Improvement Using Process Mining Technique, Doktora tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Türkiye, 2018.

37. IEEE Task Force on Process Mining, Process Mining Manifesto, Business Process Management Workshops, 169–194, 2011.

38. Dunkl R., Fröschl K., Grossmann W., and Rinderle-Ma S., Assessing medical treatment compliance based on formal process modeling, in Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society (USAB), 533–546, 2011.

39. Masic I., Miokovic M., and Muhamedagic B., Evidence Based Medicine – New Approaches and Challenges, 16 (4), 219–225, 2008.

40. Man's R. S., Aalst W. M. P. Van Der, Vanwersch R. J. B., and Moleman A. J., Process Mining in Healthcare: Data Challenges when Answering Frequently Posed Questions. International Workshop on Process-oriented Information Systems in Healthcare, 140-153, Berlin Heidelberg, Springer, 2012.

41. Senderovich A. et al., Conformance checking and performance improvement in scheduled processes: A queueing-network perspective, Inf Syst, 62, 185–206, 2016.

42. Delias P., Manolitzas P., Grigoroudis E., and Matsatsinis N., Applying process mining to the emergency department, in Encyclopedia of Business Analytics and Optimization, 168–178, 2014.

43. Ekici B., A BPMN Data Model to Keep a Multi-Perspective Process Model on the Blockchain, Master tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Türkiye, 2021.

44. Pika A., Wynn M. T., Budiono S., Hofstede A. H. M. T., Aalst W. M. P. van der, and Reijers H. A., Privacy-preserving process mining in healthcare,"Int J Environ Res Public Health, 17 (5), 2020.

45. E. Rojas et al., PALIA-ER: Bringing question-driven process mining closer to the emergency room, CEUR Workshop Proc, 1920, 1–5, 2017.

46. Zaeem R. N. and Barber K. S., The Effect of the GDPR on Privacy Policies, ACM Trans Manag Inf Syst, 12 (1), 2021.

47. Gökçay B. and Arda B., Ethical overview of health research with regard to the protection of personal health data_net Kisisel saglik verilerinin korunmasi kapsaminda saglik arastirmalarinda etik bakis, Turk Kardiyoloji Dernegi Arsivi, 47 (3), 218–227, 2019.

48. García Salvador J. and Luengo H. F., Data Preparation Basic Models, in Data Preprocessing in Data Mining, Cham: Springer International Publishing, 39–57, 2015.

49. Gökdemir A. and Çalhan A., Deep learning and machine learning based anomaly detection in internet of things environments, Journal of the Faculty of Engineering and Architecture of Gazi University, 37 (4), 1945-1956, 2022.

50. Aci M. and Doğansoy G. A., Demand forecasting for e-retail sector using machine learning and deep learning methods, Journal of the Faculty of Engineering and Architecture of Gazi University, 7 (3), 1325–1339, 2022.

51. Marin-Castro H. M. and Tello-Leal E., Event log preprocessing for process mining: A review, Applied Sciences (Switzerland), 11 (22), MDPI, N2021.

52. Bluma A. L. and Langley P., Artificial Intelligence Selection of relevant features and examples in machine, 1997.

53. Guyon I., Gunn S., Nikravesh M., and Zadeh L., Feature Extraction Foundations and Applications, 2006.

54. Zandkarimi F., Rehse J. R., Soudmand P., and Hoehle H., A generic framework for trace clustering in process mining, in Proceedings - 2020 2nd International Conference on Process Mining (ICPM), 177–184, 2020

55. Khalid S., Khalil T., and Nasreen S., A survey of feature selection and feature extraction techniques in machine learning, Proceedings of 2014 Science and Information Conference, 372–378, 2014.