

Hastalıkların Tedavi Sürecinin Protokollere Uygunluğunun Analizi: Süreç Madenciliğine Yönelik Araç Geliştirme

Kübra Akeren¹, Ayça Tarhan²

^{1,2}Bilgisayar Mühendisliği, Hacettepe Üniversitesi
Beytepe Kampüsü, 06800, Ankara, TÜRKİYE

¹kubraakeren@cs.hacettepe.edu.tr, ²atarhan@hacettepe.edu.tr

Özetçe. Veri bilimleri; (büyük) veriyi akıllıca kullanarak ürün/hizmet sunmaya yönelik süreçlerinde doğru kararlar almaya ve bu şekilde istikrar sağlamaya çalışan kurumlar sebebiyle geleceğin uzmanlık alanlarından biri haline gelmiştir. Süreç Madenciliği model tabanlı süreç analizi ile veriye yönelik analiz teknikleri arasındaki kayıp bağlantıdır ve farklı alanlarda; uygulanan sürecin keşfi, beklenen modele uygunluk kontrolü veya süreç iyileştirme amaçlarıyla kullanılmaktadır. Bu çalışmada, büyük veri içeren kayıtlar üzerinde üç temel nitelik (durum/vaka numarası-aktivite-zaman) kullanılarak analiz yapmak ve herhangi bir teşhis-tedavi sürecinin uzmana açıkça gösterilmesi sonucunda zaman, maliyet ve yaşamsal kazançlar sağlanabileceğini göstermek amacıyla, Java dilinde masaüstü uygulama olarak geliştirilmesi hedeflenen bir aracın prototipi tanıtılmıştır. Süreç Madenciliği kullanarak performans analizi için dört temel adım önerilmektedir: 1) Olması beklenen süreç modelinin (tedavi protokolü) .csv dosyası türünde alınarak araç üzerinde tanımlanması, 2) Pratikte uygulanan sürecin ilgili büyük veriden madenlenmesi, 3) Süreç uygulamalarının süreç modeline göre uygunluğunun analizi ve süreç örüntülerinin tespiti, 4) Seçilen / tanımlanan bir özellik için süreç performansının ve olası iyileştirmelerin "what-if" senaryoları yardımıyla analiz edilmesi. Prototipin işleyişi ilk üç adım için, daha önce benzer kullanımlar için hazırlanmış sağlık verisi ile örneklenmiş ve çıkarımlar raporlanmıştır.

Anahtar Kelimeler: süreç madenciliği, sağlık verisi, süreç analizi, büyük veri, uygunluk tespiti, sağlıkta süreç madenciliği.

Conformance Analysis of Disease Treatment Process To Protocols: Process Mining Oriented Tool Development

Abstract. Data science has become one of the professions of the future because organizations that are able to use (big) data in a smart way to give the right decision about their processes of creating product/services will survive by this means. Process mining is the missing link between model-based process analysis and data-oriented analysis techniques and it is used in many different fields for the discovery of process, conformance between process model and event logs, or enhancement of process. This study, intended to introduce; time, cost and vital gains are achievable by providing clearly presented process in the diagnosis and treatment of any disease to the experts; as a result of processing healthcare data with a prototype tool which is targeted as a desktop application and written in Java language, in order to make the analysis on the records that contain large data by using three basic properties (case id – activity – timestamp). At this point, four basic steps are recommended for performance analysis by using the mining process: 1) Identifying the expected process model (treatment protocol) by importing a .csv file format on the vehicle, 2) Mining the present process by using relevant big data, 3) Conformance analyzing between present models and expected models and detection of process patterns, 4) Analysing process performance and possible improvements for selected/identified property by the help of "what-if" scenarios. For the first three steps, the operation of the prototype is sampled with the healthcare data which prepared previously for similar purposes and the results are reported.

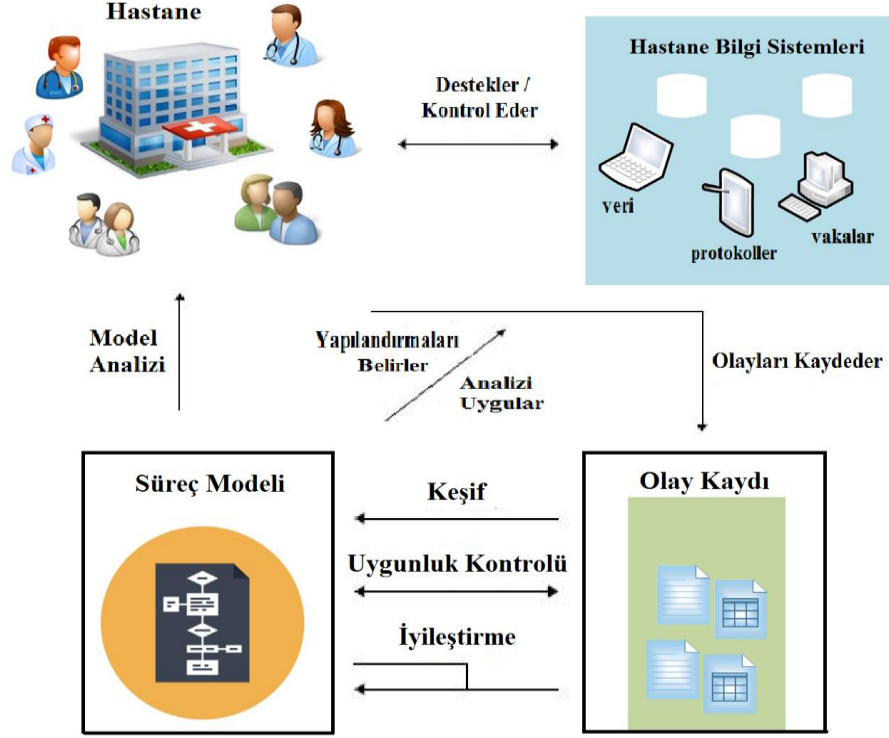
Keywords: process mining, healthcare data, process analysis, big data, conformance detection, process mining in healthcare data.

1 Giriş

Ağırlıkla insan odaklı ve bilgi yoğunluklu olan sağlık süreçleri her gün yüksek riskli hizmetleri önemli sayıda hastaya sunmaktadır^[1]. Bu nedenle, klinik kılavuzlar ve iş süreci iyileştirme, sağlık alanında önemli araştırma konularıdır. Süreç Madenciliği son on yılda iş süreci yönetimi için popüler bir teknik haline gelmiştir^[2]. Bu keşif olay günlüklerine dayalı iş süreçlerinin keşfi, doğrulanması ve geliştirilmesi dâhil olmak üzere çeşitli alanlarda uygulanmıştır^[3]. Sağlık alanında sağlık süreçleri veya onlarla ilişkili bilgi sistemlerinin değerlendirilmesini hedefleyen Süreç Madenciliği uygulamaları vardır. Ancak, bunlardan sadece birkaç tanesi uygulamalardan performans iyileştirme için nicel sonuçlar çıkardığını raporlamıştır^[4].

Süreç Madenciliği model tabanlı süreç analizi ile veriye yönelik analiz teknikleri arasındaki kayıp bağlantıdır.^[3] (Büyük) Veriyi akıllıca kullanamayan şirketler istikrar sağlamakta zorlanacağından, veri bilimlileri; geleceğin uzmanlık alanlarından biri haline gelmiştir. Veri bilimciler veri depolama ve analizine ek olarak, veri ile süreç analizini ilişkilendirmeye ihtiyaç duyar. Süreç Madenciliği olay verisi (örneğin; gözlenen davranışlar) ve süreç modelini (elle oluşturulmuş ya da otomatik keşfedilmiş) karşı karşıya getirmeyi amaçlar. Bu teknoloji yalnızca son dönemlerde erişilebilir olsa da herhangi bir tipteki işlemsel sürece (organizasyonlar ve sistemler) uygulanabilir.^[3] Tekniğin örnek uygulamaları arasında; hastane tedavi süreçleri analizi, uluslararası müşteri servis süreçlerinin iyileştirilmesi, rezervasyon sistemlerindeki müşteri göz atma davranışlarının anlaşılması, bagaj taşıma sistemi hatalarının analizi, vb yer alır. Sağlık alanı seçilmesindeki temel neden, bu alanda ilerleme kaydedilmesinin etkilerinin, diğer pek çok alandan daha önemli olduğunun düşünülmesidir.

Süreç Madenciliğinin ilk türü keşiftir. Keşif tekniği bir olay kaydı (log) alır ve hiçbir önsel bilgi kullanmadan bir süreç modeli üretir. Süreç Madenciliğinin ikinci türü uygunluk kontrolüdür. Bu aşamada var olan süreç modeli, aynı sürecin olay kaydı ile karşılaştırılır. Uygunluk kontrolü kaydedilmiş olay kaydı verilerinin doğruluğunu ve modelin doğrulanmasını, karşılıklı olarak sağlamakta kullanılabilir. Süreç Madenciliğinin üçüncü türü ise iyileştirmedir. Bu aşamada temel amaç var olan süreç modelinin, bazı olay kayıtlarından asıl süreç hakkındaki bilgiye göre genişletilmesi ya da geliştirilmesidir. Süreç modelinin performans bilgisi ile genişletilmesine örnek olarak darboğazların gösterilmesi söylenebilir.^{[3][5]} Bu çalışmada tanıtılan prototip araçta, Süreç Madenciliğinin ilk iki türü (keşif ve uygunluk kontrolü) üzerine yoğunlaşmıştır. Süreç Madenciliğinin üçüncü türü olan iyileştirmenin ise Şekil 1'de gösterildiği gibi olay kaydından madenlenen süreç modelinin (tedavi protokolü), uzmana doğru şekilde sunulması sayesinde sağlanabileceği düşünülmektedir.^[6]



Şekil 1. Sağlık Alanında Süreç Madenciliği^[6]

Bu bildirinin ikinci bölümünde konuyla ilişkili çalışmalar ve araçlar anlatılmıştır. Üçüncü bölümde Süreç Madenciliğinin sağlık verisi üzerinde koşturulabilmesini sağlayan araç prototipi tanıtılmıştır. Dördüncü bölümde benzer amaçlar için hazırlanmış radyoloji verisi üzerinde, geliştirilen prototiple bir örnekleme gerçekleştirilmiş ve gözlemlenen sonuçlar anlatılmıştır. Beşinci bölümde ise yapılan çalışmanın sonuçları, kısıtları ve gelecek çalışmalar paylaşılmıştır.

2 İlişkili Çalışmalar ve Araçlar

Bu çalışmayla ilişkili alanlarda yayımlanan çalışmalar literatür taraması şeklinde çalışmanın başlangıcında incelenmiş ve gerekli noktalarda referans olarak alınmıştır^[4]. Türkçe olarak sağlık alanında Süreç Madenciliği ile ilgili yapılmış bir çalışmaya rastlanmamıştır. Yapılan çalışmaların genelini durum analizi veya var olan araçların kullanımı ile çıkarım elde edilmesi şeklinde olduğu görülmektedir. Bazı çalışmalarda, geliştirilen araçların yazılım süreçleri, iş süreçleri gibi alanlardan elde edilen büyük veriler ile koşturulması sonucundaki gözlemler ve araçların özellikleri aktarılmıştır. İlişkili araçlar kısmında ise var olan açık-kaynaklı araçlardan erişilebilir olan araçların^[7] ^[8] özellikleri aktarılmıştır. Bu çalışmada ayrıca, örnek veriden elde edilen sonuçların doğruluğu bu araçların çıktılarıyla karşılaştırılmıştır. Aşağıda verilen Tablo 1 ve Tablo 2'de bu alanla ilgili ya da bu alanı etkileyen çalışmalar ve araçlar listelenmiştir.

Tablo 1. İlişkili Bazı Ulusal Çalışmalar

Çalışma	Temel Amacı	Sonuçları
Süreç Madenciliğinin Yazılım Süreç Doğrulama İçin Kullanımı ^[9]	Süreç Madenciliğinin yazılım süreçlerinde kullanıldığı bir örnektir.	Bu bildiri, Süreç Madenciliği tekniği kapsamında etkinlik madenciliğinin, yazılım süreç doğrulama için kullanımını anlatmaktadır. Bir kamu kurumuna ait görev yönetimi sürecinin keşfi, keşfedilen sürece göre uygulamaların doğrulanması ve uygulamadaki varyasyonların tespiti için; ProM Süreç Madenciliği aracı kullanılmıştır. Doğrulamanın başarımı istatistiksel olarak sınanmıştır. Sinama sonucunda özellikle durum esash işleyen yazılım süreçleri için, Süreç Madenciliğinin kullanılabilceği görülmüştür.
Süreç İyileştirme Çalışmaların Veri Madenciliği Yaklaşımlarının Kullanılması Üzerine Bir Çalışma ^[10]	Süreç Madenciliğinin altı sigma yaklaşımı ile birleştirilmesi amaçlanmıştır.	Bu çalışmada işletmelerin artan rekabet koşulları karşında ayakta kalabilmesi için kaçınılmaz olarak hedeflenen sürekli iyileştirme ve bu amacı benimseyen Altı Sigma yaklaşımı ve teknikleri incelenmiştir. Veri madenciliği teknikleri ile verilerin depolanmasını sağlamak ve çok büyük veri yığımlı içinde istenilen desenleri ortaya çıkarmanın hedeflendiği çalışmada, bu sayede iyileştirme yapılacak olan sürecin daha iyi anlaşılması ve analiz edilmesinin sağlanacağı belirtilmiştir.
Yazılım Geliştirme ve Test Döngüsü Üzerinde Süreç Madenciliği Yaklaşımı ^[11]	Yazılım projeleri ve iş süreçlerinde Süreç Madenciliği uygulanması amaçlanmıştır.	Bu çalışmada, SAP kullanan bir yazılım kurumundaki yazılım geliştirme ve test döngüsü olay günlükleri alınarak bir analiz yapılmaktadır. Süreç Madenciliği sayesinde karar yolları, kontrol akışı ve performans gibi yürütme bilgileri elde edilebilir. Bu bilgiler kullanılarak simülasyon modelleri üretmek mümkündür. Çalışma kapsamında yeni bir alan olarak anılan bu yaklaşıma ait araştırma ve geliştirmeye yönelik konular belirlenerek bir çalışma yapılmıştır.

Çalışmamız ile benzer nitelikler taşıyan yerel olmayan çalışmalar da bulunmaktadır. “Sağlıkta Süreç Madenciliğinin Bir Uygulaması – Hollanda Hastanesindeki Bir Durum Çalışması”^[12] isimli çalışmada Sağlık alanında Süreç Madenciliği yaklaşımı; kontrol akış, organizasyon ve performans olmak üzere üç perspektife dayandırılmıştır. Alan ve kapsam olarak bu çalışmaya en çok benzeyen çalışmadır. Aracın örnek kullanımı bölümünde kullanılan radyoloji türündeki sağlık verisi de bu çalışmada kullanılan veriden türetilmiştir. Çalışmada aynı tanı konulmuş bir hasta grubu için prosedür (bakım yolları) tanımlamak ve bölümler arası işbirliği sağlamak için Süreç Madenciliği kullanılabilceği gösterilmiştir. Sonuçları mevcut bakım akışlarının iyileştirilmesini kolaylaştıracak yeni bir anlayış sağlamak için Süreç Madenciliğinin kullanılabilceğini göstermektedir. Çalışmanın yayınladığı kaynaktan ve yakın zamanda yaptığımız çalışmadan^[4] sağlık alanında madencilik uygulanan başka çalışmalara da erişilebilir.

Çalışmamız sonucunda oluşturduğumuz aracın gereksinimlerinin ve tasarımının belirlenmesi noktasında var olan ve kullanılan diğer uygulamalardan yola çıkılarak yeni bir araç geliştirilmesi yoluna gidilmiştir. Aracın ilişkili araçlar tablosunda anlatılan diğer uygulamalardan farklı olarak temelde “ticari olmaması ve sağlık personelinin kullanması hedeflendiğinden mühendislik/teknisyenlik düzeyinde herhangi bir niteliğe ihtiyaç duyulmadan kullanılacak kolaylıkta” olması hedeflenmiştir. Tablo 2’de ayrıntıları verilen araçlardan; ProM aracı ile basitliği (kurulumu, kullanımı ve yorumlanmasının kolaylığı) yönüyle, Disco aracı ile ise ticari olmaması yönüyle farklılaştırılmıştır. Kapsamı daha dar ve odaklı olmakla birlikte geliştirdiğimiz aracın, bu iki özelliği bir arada bulundurmasıyla halihazırda kullanılan araçlara sağlık alanında kullanımda alternatif olması bekenilebilir.

Tablo 2. İlişkili Araçlar

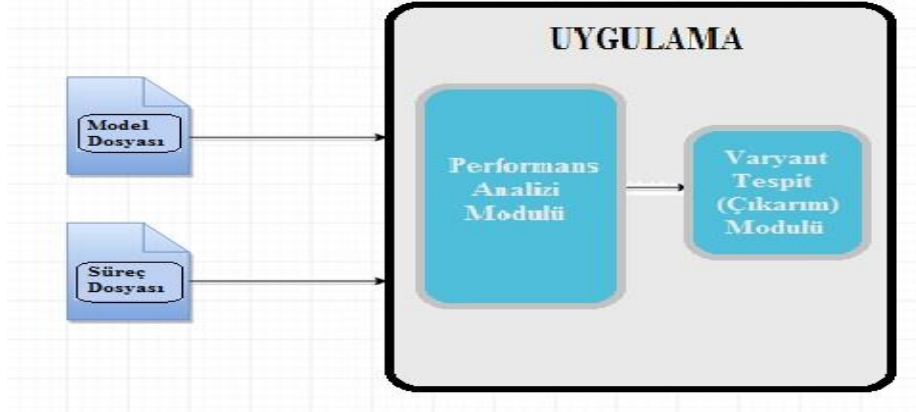
Araç	Hakkında	Özellikleri
ProM (2004) Açık-Kaynak Süreç Madenciligi Aracı ^[7]	ProM aracının geliştirilmesindeki temel neden Süreç Madenciliginde var olan araçların zayıf varsayımlara dayanması ve gerçek zamanlı Süreç Madenciligi projeleri için neredeyse hiçbir destek sağlamamasıdır. Var olan araçların bu zayıflıkları nedeniyle her yeni düşünülmüş süreç keşif tekniği için özel atanmış Süreç Madenciligi aracı üretilmesi gerekmektedir. Süreç Madenciliginin MXML formatında kullanan bu gözlem, ortamlar için "tak-çıkır" (plug-able) ProM altyapısının geliştirilmesini tetikledi ve 2004 yılından bu yana Süreç Madenciligi alanında kullanılmaya başlandı.	<ul style="list-style-type: none">- Ortak temeller için Süreç Madenciligi tekniklerinin tüm türlerini destekler.- Plug-in formunda Süreç Madenciligi tekniklerinin çeşitli türlerini destekleyen dosya dışarı aktarımını, dönüştürmeyi ve olay verisini yüklemeyi düşünmeye gerek duymadan genişletilebilir iskelete sahiptir.- Platform bağımsızdır.- Ücretsiz indirilebilir.- Akademik çalışmaları destekleyen bir geliştirici ekip ile yeni araştırmalar ve eklentilere destek sunar.- Analiz sonucunu model ve bölümlendirilmiş liste olarak göstermektedir.
Disco (2011) Ticari Süreç Madenciligi Aracı ^[8]	Disco, Süreç Madenciligi analizi için yüksek performans (büyük ve karmaşık veri setleri ile başa çıkma) ve kullanım kolaylığı odaklı olan bir tak-çıkır (bağımsız) araçtır. Bu araçtaki Süreç Madenciligi algoritması bulanık madencilik (fuzzy mining) yaklaşımına dayanır. Performans gibi diğer boyutlar da madenlenen süreç modelleri ileri görüntüleme ile analiz edilebilmektedir. Disco yayınlandığı 2011 yılından günümüze kullanılmakta olan popüler bir araçtır.	<ul style="list-style-type: none">- Basit, hızlı bir araçtır.- Ticari bir araçtır.- Geliştirime açık değildir.- Analiz sonucunu model olarak göstermektedir.- Haritacılık metaforunda kullanılan genelleme ve kusursuz soyutlamayı desteklemektedir.- Karmaşık (Spagetti) türü süreçlerde de etkili olabilmektedir.

Süreç Madenciligindeki araçlara girdilenecek verinin özel bir veri çıkarım (export) aracıyla ya da kaynak sistemlerin çeşitli çıkartım özellikleri kullanılarak elde edilmesi gerekir.^[13] Bu noktada bu çalışmanın ve oluşturulan aracın; sağlığa özel amaçlı olması (özellikle bağımsız bir uygulama olması nedeniyle verinin hastane dışına çıkartılmasını gerektirmemesi), tedavi bazında uygunluk ve performans analizine imkân tanıyor olması, yerel olarak yapılan bir çalışma olması, açık kaynak olması gibi nedenlerle diğer çalışmalardan farklı olduğu görülmektedir.

3 Prototip ve Özellikleri

Geliştirilen araç prototipi, Java dilinde masaüstü uygulama olarak yazılmıştır. Uygulama kullanıcıdan iki ayrı .csv dosyası isteyerek bu dosyalar içerisindeki (büyük) verinin algoritma sonucunda analizinin sonuçlarını kullanıcıya sunar (Şekil 2). Uygulamanın kullanımı ile uzmanın, madenlediği süreç verisi ile tedavi protokolü arasındaki farklılıkları gözlemlemesi amaçlanmıştır. Bunu takiben uzmanın, gerekli süreç iyileştirmeler için kendi alan bilgisini kullanarak yorumda bulunması beklenmektedir.

Önce seçilen (yaygın kullanılması ve dönüştürümünün kolay olması nedeniyle) .csv formatındaki sağlık verisinden okuma yapılır ve uzmanın veriyi madenlemesi sağlanır. Sonrasında madenleme sonucunda oluşan model ile aynı ekrandan seçilen tedavi protokolü verisinden oluşturulan model karşılaştırılır. Son olarak da ilgili çizgelerin örüntülerinin ekrana görsel olarak çizdirilmesi, varyant örüntülerinin sıklıkları ile beraber sıralı şekilde listelenmesi ve uzmanın yorumuna sunulması gerçekleştirilir.

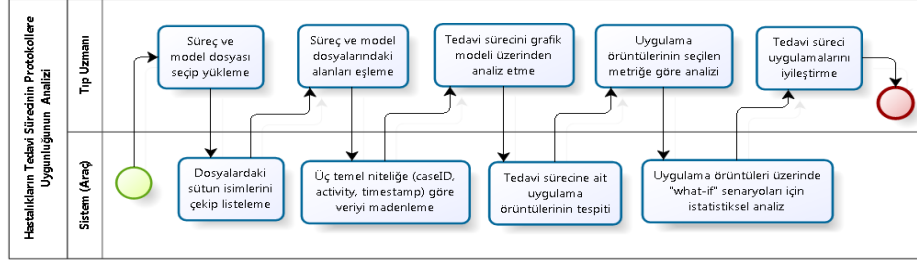


Şekil 2. Araç Tasarımı

Aracın yazılımı, model ve GKA (grafiksel kullanıcı arayüzü) olmak üzere iki ana ögede değerlendirilebilir.

Yazılımın model kısmında ilk olarak uzmandan alınan .csv dosyasının okutulması gerçekleştirilmektedir. Veri okunup ayrıştırıldığında zorunlu olan alanların eşleştirimi istenir ve hemen sonrasında bu zorunlu alanlara göre analize başlanır. Modelde veri içerisindeki her bir durum numarası için aktiviteler zaman bilgisine göre öbeklenir ve kayıt edilir. Sonrasında analiz tamamlandığında aktiviteler arasındaki geçiş için bir çizge (graph) oluşturulur. Buna uygun olarak eğer uzman belirli bir hastalık için seçimde bulunmuş ise o hastalığa ait çizge de oluşturulur. Eğer uzman veride filtreleme yapmış ise veri temizlendikten sonra analiz gerçekleştirilir. Araç, bulunan her bir "uygulama örüntüsü" (varyant) için kayıt tutmaktadır. Bu demektir ki araç, uzmanın analiz sonrası sık rastlanan varyantlar veya çok az rastlanan varyantları liste şeklinde görebilmesine olanak tanır ve böylelikle uzmanın sağlık verisindeki uygulama örüntüleri hakkında yorumlar yapması desteklenir.

Yazılımın grafiksel kullanıcı arayüzü kısmında ilk olarak, uzmanın 'analiz edilecek olan süreç verisini' ve 'süreç verisi ile karşılaştırılacak olan gerçek-hayat prosedürünü içeren model verisini' dosya seçme tuşuna basarak seçmesi için gerekli ekran bulunmaktadır. Bu ekranda dosyalar seçildikten sonra dosyalara ait olan sütun isimleri listelenir. Uzman, analiz verisi ile model verisindeki sütun adları arasında eşleme yaparak; araca, hangi niteliğin veride hangi sütunda yer aldığı bilgisini girmiş olur. Sonrasında eşleme yapılan sütunlarda genel analiz başlatılır. Eğer uzman yalnızca belirli bir hastalık için analiz yapmak isterse ikinci ekran açılır ve veride yer alan tüm hastalıklar uzmana bu ekranda listelenir. İlgili hastalık seçilerek yine tüm veride yapıldığı şekilde analiz başlatılır ve sonuçları ekrana yansıtılır. Diğer bir ekranda ise bulunan varyantlar verideki görülme sıklıkları bilgisine göre çoktan aza listelenmektedir (Şekil 3).



Şekil 3. Araç Aktivite Akış Diyagramı

Analizin başlangıcında uzman, süreç için sisteme tanımlayarak kullandığı .csv uzantılı dosyayı uygulama içinde bir kez tanıtarak aynı dosya üzerinden tekrar ayrıntılı süreçleri başlatabilir. Seçilen hastalıkla ilgili kayıtlar görüntülenebildiği gibi bu kayıtlara ait grafiklerin de görsel olarak çizimi oluşturularak kullanıcı dostu bir yapıda olması amaçlanmaktadır. Bu nedenle grafiklerin görsel çizimi üzerindeki çalışmalar devam etmektedir. Araç Java8 üzerinde geliştirildiği için herhangi bir işletim sistemine bağlı olmadan ilgili JDK'nın mevcut olduğu her bilgisayarda kolaylıkla çalıştırılabilir. Ayrıca .jar dosyası olması sebebiyle kurulum gerektirmeden çalıştırılabilir.

Tablo 3. Araç ve Çalışmada Kullanılan Veri ile Alakalı Terimlerin Tanımlamaları

Case ID	Hasta (vaka) numarasıdır. Kayıt altına alınan her bir hasta için farklıdır.
Activity	Uzman tarafından gerçekleştirilip cihazlarca kayıt altına alınan eylemdir. Örneğin buradaki radyoloji verisi için ct abdomen (karın bilgisayarlı tomografisi) bir aktivitedir.
Timestamp	Her bir aktivitenin gerçekleşme zaman bilgisidir.
Diagnosis	Her bir hastalığın ismidir.
Protokol	Tıp uzmanının gerçekte herhangi bir hastalıkta teşhis ve tedavi için uyguladığı sıralı işlemler bütünüdür. Örneğin buradaki radyoloji verisi için maligniteit endometriyum (rahim içi kanseri) hastalığında; önce echo onderbuik (karın boşluğu echo'su) sonra ct abdomen (karın bilgisayarlı tomografisi) yapılması gerektiğinin bilgisini içerir. Bu çalışmada uzmanın tıpkı analiz edilecek veri gibi protokol verisini de .csv dosyası şeklinde sisteme tanıtmayı düşünülmemektedir. Şekil 2'deki ve Şekil 3'teki Model dosyası bu dosyadır.
Süreç	Tıp uzmanının analiz etmek istediği büyük veriyi içeren dosyadır. Bu dosyada herhangi bir cihaz üzerinden alınan sağlık verisinin ilgili bilgileri yer almaktadır. Bu çalışmada bu dosyanın .csv dosyası şeklinde tıp uzmanı tarafından sisteme ekrandan seçilerek yüklenmesi beklenmektedir.
Varyant (Uygulama Örüntüsü)	Herhangi bir hastalık için görülen çeşitli yolların listesidir. Burada amaçlanan tıp uzmanının süreç verisinden seçtiği herhangi bir hastalık için tüm vakalarda (Case ID) nasıl yöntemler izlendiğini görebilmesi ve sıklık sayıları sayesinde en sık-en seyrek rastlanan yöntemleri de ayrıca inceleyebilmesi ve analiz edebilmesidir.

4 Prototipin Örnek Kullanımı

Bu bölümde çalışmada geliştirilen aracın Flemenkçe (Dutch) dilinde benzer bir çalışmadan elde edilip daraltılan bir sağlık verisi^[14] ile koşutlanması örneklenecektir. Bu veri daha büyük bir veriden yalnızca org:group bilgisi Radiology (Radyoloji)* olarak seçilmiş bir örnek kümesidir. Araç ve veri ile alakalı terimler Tablo 3'te verilmiştir.

Örneklemede kullanılan veri hakkındaki bilgiler Tablo 4’de verilmiştir.

Tablo 4. Örnek Veri^[14] Bilgileri

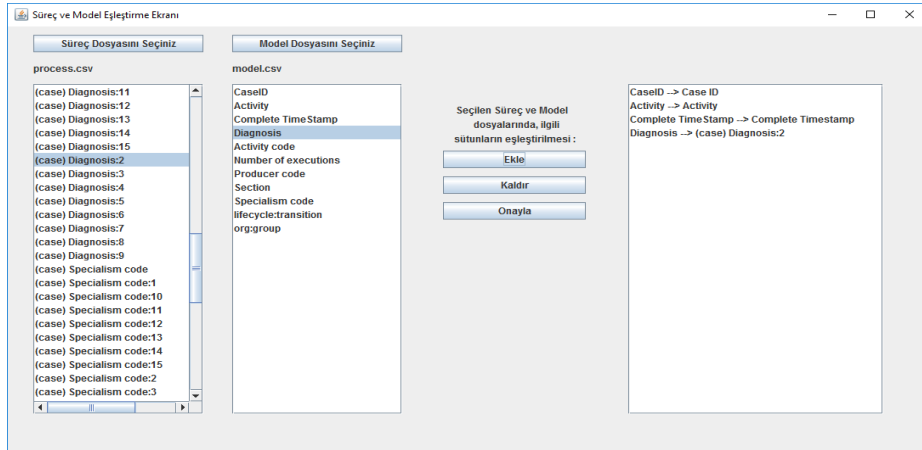
Satır Sayısı :	3172
Sütun Sayısı :	97
Sütunların İçerdiği Bilgiler	Case ID, Acvitiy, Complete Timestamp, Variant, (case) Age, (case) Diagnosis, (case) Diagnosis Combination ID, (case) Diagnosis code, (case) Specialism code, (case) Treatment code, Activity Code, Number of Executions, Producer code, Section, Specialism code, lifecycle:transition, org;group
Etkinlik Sayısı :	124

*(Radyoloji, x ışınları ve diğer görüntüleme yöntemlerinin tıpta tanı ve tedavi amacıyla kullanılmasıdır. Tanı ve tedavi amacıyla kullanılan yöntemlerden bazıları; ultrason, bilgisayarlı tomografi (CT), manyetik rezonans görüntüleme(MR), mamografi, floroskopi ve X ışını kullanan diğer bazı yöntemler olarak sıralanabilir.)^[15]

Araç tek başına çalışmaktadır ve başlangıç olarak uzmandan iki adet .csv dosyası olarak işleme başlar. Bu dosyalardan biri uzmanın incelediği alanda gerçek hayatta uygulanması beklenen tedavi protokolünden bir model oluşturulması amacı ile alınır. Alınan diğer .csv dosyası ise uzmanın asıl inceleyeceği süreci barındıran dosyadır. Süreç ve Model Eşleştirme Ekranında uzmandan ilgili dosyalar alındıktan sonra süreç dosyasında madenleme için gerekli alanların karşılıklarının işaretlenmesi için bir ekran sunulur. Bu ekran üzerinden uzman madenlemek istediği alanları işaretler (Şekil 4).

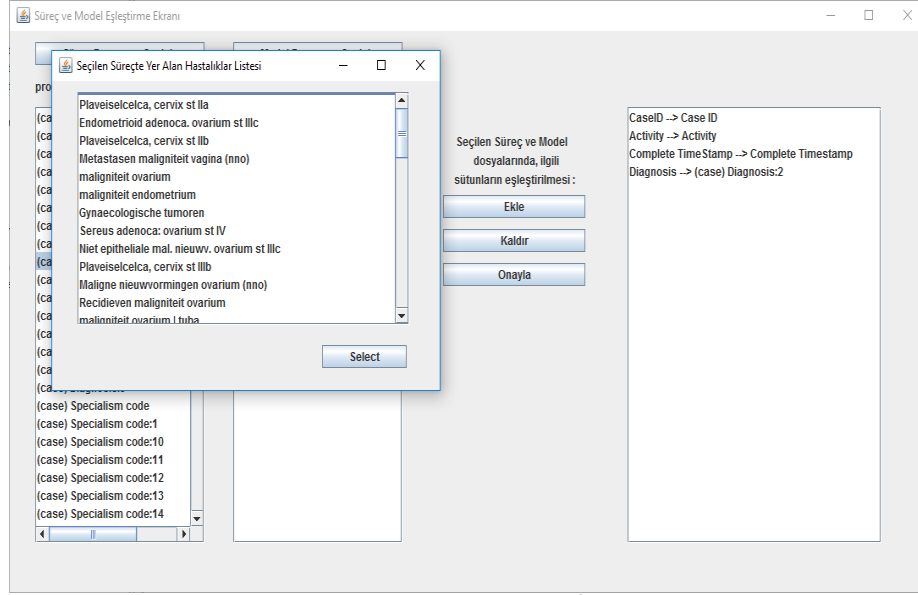
Bu çalışmada geliştirilen araç dört temel ekrana sahiptir; “Süreç ve Model Eşleştirme Ekranı”, “Hastalık Seçme Ekranı”, “Süreç Modeli Ekranı”, “Uygulama Örüntüsü Listeleme Ekranı”. Bu ekranlara ilişkin tanımlar ve örnek görünüm izleyen paragraflarda verilmiştir.

Süreç ve Model Eşleştirme Ekranı (Şekil 4): Uzman Süreç ve Model dosyasını seçerek bu iki dosyada ilgili ekranların eşleştirilmesini sağlar. Bu örneklemede ilgili veri süreç dosyası olarak, ilgili veriden elde edilecek analize odaklı özelliğe uygun daha küçük bir veri ise model (tedavi protokolü) dosyası olarak kullanılmıştır.



Şekil 4. Süreç ve Model (Tedavi Protokolü) Arasında Madenleme İçin Gerekli Özellikleri Eşleştirme

Hastalık Seçme Ekranı (Şekil 5): Uzman madenleme işlemini ilgili büyük verinin tamamı yerine veride tespit edilen tek bir hastalık üzerinden yapmak isterse bu ekrandaki listede seçilen verideki hastalıklar listenecek ve uzman bu listeden herhangi bir hastalık seçerek yalnızca ilgili hastalığa ait modeli oluşturabilecektir.

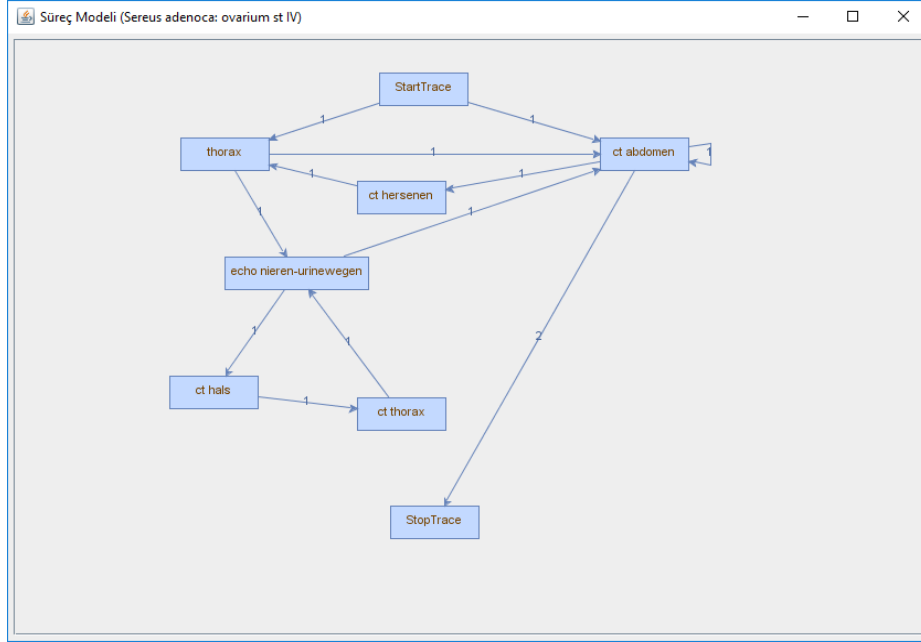


Şekil 5. Süreçte Yer Alan Hastalıklar Listesi İçerisinden Seçim Yapma

Süreç Modeli Ekranı (Şekil 6): Önceki ekranda seçilmiş olan hastalık ile ilgili bir grafiksel modelin oluşturulması sağlanır. Bu çizgenin okuması örnek olarak şu şekilde yapılabilir*:

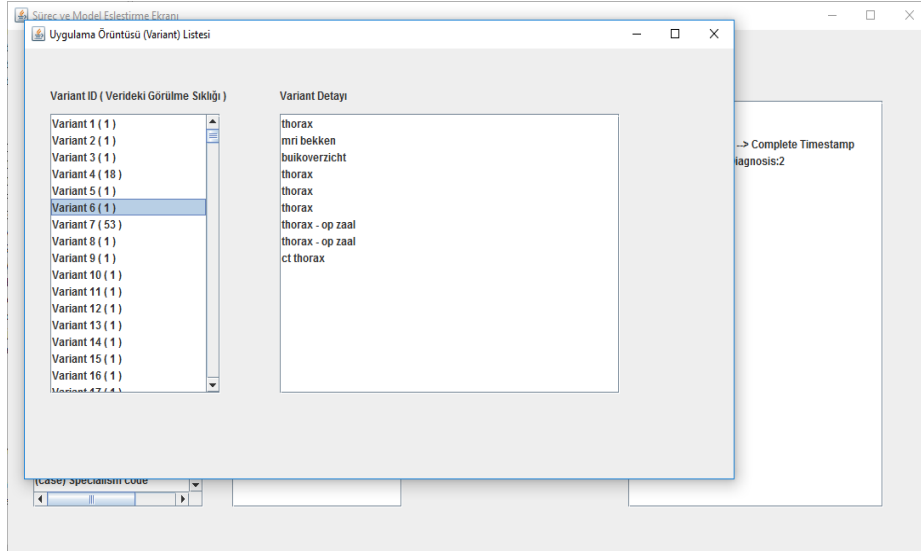
“Bu uygulama örüntüsüne ait süreç kayıtlarında (hastalık StartTrace’den yalnızca iki kez başlatıldığından), hastalık yalnızca iki vakada (iki farklı Case ID’de) görülmüştür. Başlangıç izi olarak bir kez göğüs filmi (thorax) ve bir kez de karın bilgisayarlı tomografisi (ct abdomen) görülmüştür. Vakalar göğüs filmi ile başladığında ardından karın bilgisayarlı tomografisi ya da böbrek üriner sistem echo’su (echo nieren-urinewegen), karın bilgisayarlı tomografisi ile başladığında tekrar karın bilgisayarlı tomografisi ya da beyin bilgisayarlı tomografisi (ct hersenen) görülmüştür. Göğüs filmi görülen vakada ardından Böbrek üriner sistem echo’su görülmüş ise ardından karın bilgisayarlı tomografisi ya da boyun bilgisayarlı tomografisi (ct hals) görülmüştür. ...”

*(Kullanılan radyoloji verisi Flemenkçe dilinde olduğu için şekilde görülen olay kayıtlarına ait çevirilerin tamamı parantez içerisinde belirtilmiştir.)



Şekil 6. Seçilen Hastalığa Ait Sürecin Grafiksel Olarak Modelinin Çizdirilmesi

Uygulama Örüntüsü Listeleme Ekranı (Şekil 7): Uzmana ilgili büyük veri üzerinden madenlenerek çıkarılan tüm olası süreç örüntülerini dosyada bulunma sıklıklarına göre listelenebilmesi ve her bir örüntünün modelinin tek tek görüntülenebilmesi sağlanır.



Şekil 7. Sürece Ait Uygulama Örüntülerinin (Varyant) Listelenmesi

5 Sonuç, Kısıtlar ve Gelecek Çalışmalar

Bu çalışmada, Süreç Madenciliğinin sağlık alanında temin edilen bir veri üzerinde uygulanabilirliğinden örneklem yapılmış, bu amaçla geliştirilen aracın prototipi tanıtılmış, benzer çalışmalar ve araçlar hakkında açıklamalar verilmiştir. Süreç Madenciliğinin neden önemli olduğu ve sağlık alanında kullanımının olası faydaları tartışılmıştır. Yapılan örnekleme için bir radyoloji uzmanından geribildirim alınmıştır ve dönüt, tekniğin sağlık alanındaki analizlerde fayda sağlayabileceği yönündedir. Ne var ki aracın kullanılabilirliğine ve işe yararlığına ilişkin deneysel çalışmalara ihtiyaç vardır ve gelecek çalışmalar dâhilinde planlanmaktadır.

Süreç Madenciliği özellikle önemli operasyonel süreçleri olan organizasyonlar için önemli bir araçtır. Bir taraftan bilgi sistemlerinde kaydedilen verilerdeki artış oldukça fazladır. Diğer yandan da; süreçler ve içerdikleri bilgiler; uyum, verimlilik ve müşteri memnuniyeti gibi sebepler nedeniyle mükemmelleştirilmelidir. Süreç Madenciliği oldukça kolay uygulanabilir. Ancak belirtilmelidir ki; Süreç Madenciliği gelişmekte olan bir disiplin olduğundan, çeşitli zorlukları ve kısıtları vardır. Bu zorluklar; olay verilerinin bulunması, ayarlanması ve temizlenmesi, farklı karakteri olan karmaşık olay logları ile başa çıkmak, temsil noktalarını yaratmak, kavramsal sapmalar ile başa çıkmak, vs. olarak sıralanabilir.^[5] Bu çalışmada; veri entegrasyonu, etik kısıtlar ve verinin tümleşik olmaması temel kısıtlar olarak düşünülebilir.

Geliştirilen aracın gerek yeni sayılabilecek bir alanda yapılması gerekse gelişen teknolojik ve tıbbi imkânların hızlı artışı sayesinde gelecek çalışmalara örnek teşkil edeceği düşünülmektedir.^{[16][17]} Özellikle sağlık alanındaki sistemler süreç-haberdar hale dönüştürülüp sağlık alanında veri kaydı için standartlar geliştirilebilirse Sağlıkta Süreç Madenciliği daha anlamlı hale gelebilir.^[18] Bu çalışmada geliştirilen araç, sağlık verisinin etik nedenlerle kurum dışına çıkartılamaması durumunda yerinde analiz edebilme imkânı tanımak üzere hedeflenmiştir. Çalışma sonucunda elde edilenlerin ileride başka alanlarda kullanılması ya da bu alanda daha derinlemesine yapılacak çalışmalara kaynak oluşturması beklenmektedir. Bu bağlamda araç üzerinde grafiksel ve algoritmik iyileştirmeler sürmektedir.

Referanslar

- [1] Brender, J., Nöhr, C., McNair, P.: “Research needs and priorities in health informatics”. Int. Journal of Medical Informatics 58–59, pp. 257–289 (2000).
- [2] Dumas, M., Rosa, M.L., Mendling, J., Reijers, H.A.: Fundamentals of Business Process Management. Springer (2013).
- [3] W.M.P. van der Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes. DOI 10.1007/978-3-642-19345-3_1, © Springer-Verlag Berlin Heidelberg 2011.
- [4] Erdoğan, T., Tarhan, A.: “Process Mining for Healthcare Process Analytics”, in proceedings of IWSM Mensura 2016, 5-7 Ekim 2016, Berlin. (*kabul edildi*)
- [5] IEEE Task Force on Process Mining, Process Mining Manifesto (Süreç Madenciliği Manifestosu-Türkçeye çeviren: Aylin Çahin), in BPM Workshops, volume 99 of Lecture Notes in Business Information Processing. SpringerVerlag, Berlin, 2011, 169–194.
- [6] Eric Rojas, Jorge Munoz-Gama, Marcos Sepúlveda, Daniel Capurro, “Process Mining in Healthcare: A literature Review” Journal of Biomedical Informatics 61 (2016) 224–236
- [7] ProM(2004) Açık-Kaynak Süreç Madenciliği Aracı; <http://www.promtools.org/doku.php>
- [8] Disco(2011) Ticari Süreç Madenciliği Aracı; <https://fluxicon.com/disco/>
- [9] Gürgen, Tuğba ve Tarhan, Ayça, (2011). Süreç Madenciliğinin Yazılım Süreç Doğrulama İçin Kullanımı; <http://www.uyms.org.tr/2011/bildiriler/b37.pdf>
- [10] Yalçın Pirinçliler, Esin Cumhuri ve Şen, Ali, (2012). Süreç İyileştirme Çalışmalarının Veri Madenciliği Yaklaşımlarının Kullanılması Üzerine Bir Çalışma; <http://www.sobbiad.mu.edu.tr/index.php/asd/article/viewFile/265/395>
- [11] Saylam, Rabia ve Sahingoz, Ozgur Koray, (2014). Yazılım Geliştirme ve Test Döngüsü Üzerinde Süreç Madenciliği Yaklaşımı; http://ceur-ws.org/Vol-1221/62_Bildiri.pdf
- [12] R. S. Mans, M. H. Schonenberg, M. Song, W. M. P. van der Aalst, P. J. M. Bakker, “Application of Process Mining in Healthcare – A Case Study in a Dutch Hospital”, Biomedical Engineering Systems and Technologies Communications in Computer and Information Science Volume 25, 2009, pp 425-438, Springer Berlin Heidelberg.
- [13] https://www.wiso.uni-hamburg.de/fileadmin/wiso_fs_wi/Publikationen/Michael/Gehrke_und_Werner_-_2013_-_Process_Mining_Pre-print_Version.pdf
- [14] Prototip İçin Kullanılan Verinin Kaynağı; <http://data.4tu.nl/repository/uuid:d9769f3d-0ab0-4fb8-803b-0d1120ffc54>
- [15] <https://tr.wikipedia.org/wiki/Radyoloji>
- [16] C. Fernandez-Llatas, T. Meneu, J. Miguel Benedi, and V. Traver, “Continuous Clinical Pathways Evaluation By Using Automatic Learning Algorithms,” in Healthinf 2011: Proceedings of the International Conference on Health Informatics, 2011, pp. 228–234.
- [17] T.G. Erdogan, A. Tarhan and N.A. Karagoz. An Integrated Infrastructure Using Process mining Techniques for Software Process Verification, Uncovering Essential Software Artifacts through Business Process Archaeology, Idea Group Inc., 2013. (DOI: 10.4018/978-1-4666-4667-4).
- [18] Sağlıkta Süreç Madenciliği; <http://www.processmining.org/health/start>