ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

VERİ MADENCILIĞI TEKNİKLERİ KULLANARAK SUÇ ANALİZİ VE ÖNLEME

Sientchanwa SORO

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

ANKARA
2020

ÖZET

Yüksek lisans Tezi

VERİ MADENCİLİĞİ TEKNİKLERİ KULLANARAK SUÇ ANALİZİ VE ÖNLEME

Sientchanwa SORO

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman : Pr.Dr. Iman ASKERBEYLI
Eşdanışman : Dr. Öğr. Üyesi Bülent TUĞRUL

Son yıllarda, terörizmin artması ve bu nedenle, küresel güvenlik için artan endişe ile, suçu analiz etmek ve önlemek için kullanılabilecek veri madenciliği gibi etkili tehdit algılama sistemlerinin kurulması zorunlu hale gelmiştir. Potansiyel olarak tehlikeli durumları tanımlamak ve yetkililerin uygun önlemleri almaları konusunda uyarılması gerekmektedir. Suç analizi ve önlenmesinin önemli bir unsuru, polis ve diğer ceza adaleti kurumları tarafından toplanan büyük bilgi veritabanlarının sorgulanmasıdır. Büyük verilerden gizli bilgileri keşfetme süreci olan veri madenciliği, suçun araştırılması, analizi ve önlenmesi için önemli bir araçtır. Uzman ve polise, kalıpları ve eğilimleri keşfetmeleri, suç ve suçlu arasındaki ilişkileri bulmaları için yardımcı olacaktır. Veri madenciliğinde kümeleme ve sınıflandırma gibi belirli teknikler kullanılabilir. Bu yaklaşımlar araştırmacıların büyük veritabanlarını etkili bir şekilde analiz etmelerine yardımcı olacaktır. Yaklaşımımız toplumun iyileştirilmesine katkıda bulunur. Soruşturma kurumlarına suç tespitinde yardımcı olmak ve suçluların kimlik tespiti ve böylece suç oranlarının azaltılması hedeflenmektedir.

Ağustos 2020, 56 sayfa

Anahtar Kelimeler : Veri Madenciliği, Kümeleme, Sınıflandırma, Suç Analizi ve Önleme, Düşük uygulama.

ABSTRACT

Master Thesis

CRIME ANALYSIS AND PREVENTIONS USING DATA MINING TECHNIQUES

Sientchanwa SORO

Ankara University
Graduate school of Natural and Applied Science
Department of computer Engineering

Supervisor: Pr.Dr. Iman ASKERBEYLI
Co- Supervisor: Asst.Prof. Dr. Bülent TUĞRUL

In recent years, with the rise of terrorism and, therefore, the growing concern for global security, it has become essential to set up effective threat detection systems, such as data mining that can be used to analyze and prevent crime, but also able to identify and recognize potentially dangerous situations and alert the authorities to take appropriate measures. An essential element of crime analysis and prevention is the interrogation of large databases of information collected by the police and other criminal justice agencies. Data mining, the process of discovering hidden information from big data, is an important tool for investigating, analyzing and preventing crime. It will help the specialist and the police to discover patterns and trends, to find relationships between crime and criminal. Certain techniques in data mining can be used, such as clustering and classification. These approaches will help investigators to effectively analyze large databases. Our approach contributes in the betterment of the society by helping the investigating agencies in crime detection and criminals' identification, and thus reducing the crime rates.

August 2020, 56 pages

Keywords: Data Mining, Clustering, Classification, Crime Analysis and Prevention, Law-enforcement.

TE ÜR

Çalışmalarımı destekleyen, bilgi ve tecrübeleriyle yönlendiren, bu çalısmanim ortaya çıkarılmasında çok fazla emeği olan hocam sayın Prof. Dr. Iman ASKERBEYLI (Ankara Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı); çalismalarim süresince her türlü fedakarlığı esirgemeyen değerli babam Ernest Eadelman'a ve değerli çalışma arkadaşlarıma en derin duygularla teşekkür ederim.

Sientchanwa SORO
Ankara, Ağustos 2020

İÇİNDEKİLER

ŞEKİLLER DİZİNİ

ÇİZELGELER DİZİNİ

1. GİRİŞ

Suç insanlığa zararlıdır, insana karşı bir suçtur ve aynı zamanda sosyal sıkıntıdır ve toplumumuza çeşitli şekillerde mal olur. ABD'de 5.000'den fazla kurbanı olan 11 Eylül'ü hatırlıyoruz. Yoksulluk, eşitsizlik, suçla ilgili sorunları ve yüksek kentleşme oranlarını, politik geçişleri, kentsel yoğunluğu, nüfus artışını ve tasarım eksikliğini, yönetimi ve uygun kentsel planlamayı da içeren bir dizi faktör. Suç analizi günümüzde kolluk kuvvetleri tarafından suçun önlenmesinde en önemli alanlardan biridir**.** Suçu analiz yoluyla değerlendirmek de suç önleme çabalarına katkıda bulunur. Bununla birlikte, suçtaki kalıpları ve eğilimleri bulmak ağır bir iştir. Bir paterni tanımlamak için, veri analisti belirli bir suçun bilinen bir paterne sığması çok zaman alabilir. Suçları takip etmek için yeni teknolojilerin artan kullanımı ile bilgisayar veri analistleri, kolluk kuvvetlerine ve diğer ceza adalet kurumlarına suçları çözmeyi hızlandırmak için ve aynı zamanda önlemek için hayati yardım sağlamaktadır. Bu araçlar, büyük veri toplanmasına dayanmaktadır. Bu veri toplama, Weka gibi yazılım araçları yardımıyla gerçekleştirilecektir. Toplanan veri miktarı ile bunları yönetmek çok zorlaşır.

Veri Madenciliği veya Veritabanlarında Bilgi keşfi, daha önce bilinmeyen ve potansiyel olarak yararlı bilgilerin örtülü olarak önemsiz bir şekilde çıkarılmasıdır. Gizli bilgilerin, beklenmedik kalıpların ve büyük veritabanlarından yeni kuralların keşfedilmesi ile ilgilenir. Veri Madenciliği, verilerde geçerli, potansiyel olarak, nihayetinde anlaşılabilir bir yapıyı tanımlama işlemidir. Suç veri kümelerinin yüksek hacmi ve bu tür veriler arasındaki ilişkinin karmaşıklığı, kriminolojiyi veri madenciliği tekniklerini uygulamak için uygun bir alan haline getirmiştir. Veri madenciliği yaklaşımları, suçların çözümünde kolluk kuvvetlerine büyük bir destek sağlar, suçlar ve profil oluşturma çamurluklar arasında bağlantılar bulur, suç ağını belirler, suçları eşleştirir, şüpheli üretir ve suç faaliyetlerini önler. Veri madenciliği, polis tarafından toplanan verileri tahmin etmek ve analiz etmek için Kümeleme, Sınıflandırma ve İlişkilendirme Kuralları da dahil olmak üzere birçok görev ve teknik içerir.

2. İLGİLİ ÇALIŞMA

Cezai fiiller, basit bir hızlı suçtan kitle öldürmeye veya terörist bir eyleme kadar değişebilir. Bu nedenle, suçlar tehlikelerine göre kategorilere ayrılabilir ve yasa ile farklı şekilde cezalandırılabilir. Bu suçlar polis ve diğer güvenlik kurumları tarafından veya özel dedektiflerle işbirliği içinde analiz edilir.

Çizelge 2.1 Suç Türleri ve Farklı Kolluk Güçleri

HUKUK İCRA FARKLI SEVİYELERİNDE SUÇ TÜRLERİ	
Suç türü	**Yerel kolluk kuvvetleri**
Trafik Kuralı Ihlalleri	Hız, pervasız sürüş, çarpışmada maddi hasara veya kişisel yaralanmaya neden olma, uyuşturucu veya alkolün etkisi altında sürüş, vur-kaç, "yol öfkesi"
Seks Suçu	Cinsel istismar, tecavüz, cinsel saldırı, çocuk tacizi, çocuk pornografisi, fuhuş
Hırsızlık	Soygun, hırsızlık, hırsızlık, motorlu taşıt hırsızlığı
Dolandırıcılık	Kara para aklama, sahtecilik, sigorta sahtekarlığı, yolsuzluk ve rüşvet, varlıkların kötüye kullanılması
Kundakçılık	Depo veya apartman gibi mülke zarar vermek için kasıtlı olarak yangın çıkarmak
Çete / Uyuşturucu Suçları	Yasadışı uyuşturucu bulundurmak, dağıtmak ve satmak
Şiddetli Suç	Cinayet, ağır saldırı, silahlı soygun, zorla tecavüz, nefret suçu
Siber Suç	Kredi kartı ve avans ücreti sahtekarlığı, hileli Web siteleri ve yasa dışı çevrimiçi kumar ve ticaret gibi İnternet sahtekarlığı ; ağa izinsiz giriş ve bilgisayar korsanlığı ; virüs yayılması ; siber korsanlık ve siberterörizm ; çocuk pornografisi dağıtımı ; kimlik Hırsızı

2.1 Suç Analizinin Tanımı

Suç analizi, 1960 ve 1970'lerde organize suçların büyümesiyle ortaya çıkmıştır. Örneğin, hipotezler geliştirmek, gerçekleri yeniden yapılandırmak, suçların aynı fail tarafından işlenip işlenmediğini belirlemek, suç ağlarının işleyişini anlamak ve suç faaliyetinin kapsamını ve özelliklerini araştırmak için bir dizi standart analitik yöntem geliştirildi. Bu düşünceye yol açan ilk faktör aşağıdaki gibidir :

Grossrieder ve Ribaux, (2013) suç soruşturmalarının artan karmaşıklığı, çeşitli türlerdeki etkileyici miktarda bilgiyi işlemek için yapılandırılmış ve daha homojen bir çalışma yöntemi aramasını gerektirmiştir. Seri katillerin seyrini veya diğer örnekleri, finansal düzenlemeleri veya uyuşturucu kaçakçılığı kanallarını referans alabiliriz. İkinci faktör bilgisayar teknolojilerinin geliştirilmesidir. Bilgisayar istediği kadar veri işleyebilir. Şimdi sömürülmesi ve kesilmesi mümkün olan veri hacimleri etkileyici. Gerçekten de, bilgisayar yazılımı, binlerce bilgiyi veritabanı ve dosya öğelerinden etkili bir şekilde yönetebilir. Bir araştırmanın farklı unsurları arasındaki bağlantıların tanımlanması, bu teknolojik araçların potansiyeli ile kolaylaştırılmıştır. Buna ek olarak, ceza analizi yazılımı birçok şekilde (diyagramlar, kronolojik ve / veya mekansal-temporal diyagramlar) yapılan işleri gerçekleştirme imkanı sunar. Suç analizi, operasyonel ve idari personele suçun korunması için kaynak dağıtımının planlanmasında yardımcı olmak için suç modelleri ve eğilim korelasyonları hakkında zamanında ve ilgili bilgi sağlamayı amaçlayan bir dizi sistematik analitik süreç olarak tanımlanmaktadır. Cezai faaliyetlerin önlenmesi ve bastırılması, soruşturma sürecine yardımcı olarak ve endişeleri ve davaların çözümünü artırarak. Devriye konuşlandırması, özel operasyonlar ve taktik birimleri, soruşturmalar, planlama ve araştırma, suç önleme ve idari hizmetler dahil olmak üzere bir dizi hizmet işlevini destekler.

Suç analizi, endekslerin toplanması ve yapılandırılması sürecini desteklemek için genel bir yöntemle yapılandırılmış teknikler önererek adli soruşturmanın bir parçasıdır. Suçun önlenmesi alanında, suçun azaltılmasında ve toplum güvenliğinin güçlendirilmesinde etkili olduğu kanıtlanmış programları desteklemek için, kanıta dayalı yaklaşım birkaç yıldır kullanılmaktadır. Kanıta dayalı yaklaşım, titiz bir araştırma ve değerlendirme metodolojisi ile etkili olduğu kanıtlanmış ve sonuçları tutarlı bir pozitif eğilim gösteren program ve uygulamaları ifade eder. Çalışan etkili uygulamaların belirlenmesi ve kanıta

dayalı yaklaşım bilgisinin program geliştirmeye uygulanması, suç önleme ve düzeltmeleri, fiziksel ve zihinsel sağlık hizmetleri ve eğitim gibi çeşitli alanlarda en iyi uygulamaların kullanılmasını sağlamak için gereklidir. Suçlar genellikle sözde kolluk kuvvetlerinde bir suç analizi olarak varlıklarını haklı çıkarma eğilimindedir. Suçu analiz etmenin mantıklı olmasının bazı nedenleri. Bazı örnekler aşağıda listelenmiştir.

- Güvenliği güçlendirmek için kolluk görevlilerini bilgilendirmek amacıyla suçun analizi ve önlenmesi.
- Kolluk kuvvetlerinde, ceza adalet sisteminde ve kamu malı olan zengin bilgi birikiminden yararlanmak için suçu analiz edin.
- Suç verilerinin analizi ve veri önleme tekniklerinin kullanımı Suçun önlenmesi için teknikler.
- Kolluk kuvvetleri tarafından sınırlı kolluk kaynaklarının kullanımını azaltmak için suçu analiz edin.

Suç analizi, ilişkileri saptamak, ezberlemek ve temsil etmek için bilgi yönetimi tekniklerini kullanan sistematik bir yöntem olarak kabul edilir. Bilgilerin yapılandırılmış kullanımı yoluyla, akışları kontrol etmek ve genel bakışı korumak için erişilebilir bilgileri yönetmeye yardımcı olan, akıl yürütmeyi yapılandırmayı ve özellikle önyargıyı en aza indirmeyi amaçlayan ve nihayetinde bilgiyi perspektife koyan genel bir süreci açıklar. Karar vericiler. Genellikle, suç analizi görevleri polis veya soruşturma ekibinin birlikte çalışması için sıkıcı bir süreç olabilir. Suç mahallinden ayrılırken suçlular, suçluları tanımlamak için bir ipucu olarak kullanılabilecek bazı izler bırakıyor.

Boba, R. (2001) suç dizisi ve bir suç işlerken birkaç suçlunun izlediği kalıplar, suçun analizini kolaylaştırır. Bu süreç, suçluları tanımlamak ve sadece yerel halk tarafından verilen ipucuna veya bilgilere dayanarak daha fazla bilgi almak için izlenecek birkaç prosedür içermektedir. Suçlu, önceki suç şekillerine karşı test edilen suç mahallindeki bilgilere dayanarak analiz edilebilir ve tahmin sonuçlarını etkileyebilecek bilgileri test etmek ve ilerlemek için ima edilen yöntemle değerlendirilebilir. Öngörme ayrıca suçları önceden tespit etmek için veya sistem tarafından tanımlanan hassas alanlara daha fazla polis eklenerek faydalı olabilir. Polis karakolları, önceden suç işleme olasılığı olduğunda özel kuvvet uygulayabilir. Bu tür bir sistem vatandaşlar arasında barış ve refah olmasını sağlayacaktır.

2.2 Suç Analizi Türleri

2.2.1 Taktik Suç Analizi

Operasyon personelinin (devriye ve araştırmacılar) eğilimleri, kalıpları, serileri, krizleri ve özel ve acil suç noktalarını belirlemelerine yardımcı olmak, soruşturma için yollar sağlamak ve davaları çözmek için bilgi sağlayan analitik bir süreç. Analiz, suç faaliyetlerinin suç yöntemi, zaman, tarih, yer, şüpheli, araç ve diğer bilgi türleriyle ilişkilendirilmesini içerir. Taktiksel suç analizi, memurlara ve araştırmacılara belirli ve acil suç sorunlarını belirleme ve anlama konusunda yardımcı olmak amacıyla suçların nerede, ne zaman ve nasıl gerçekleştiği hakkında bilgi geliştirmek için verilerin analiz edilmesini içerir. Taktiksel analizin amacı, halihazırda meydana gelen bir suç sorununa hızlı bir tepki verilmesini teşvik etmektir. Taktik suç analiz birimleri, devriye memurları ve araştırmacılarla yakın çalışacaktır. Taktik bir suç analizi olarak rollerden biri, mevcut suçlu kalıplarını tespit etmektir, gelecekteki olası suç olaylarını öngörme faaliyetidir. Taktik suç analizinin üç ana amacı vardır : suç olaylarını birbirine bağlayarak suç türlerini ve sıralarını belirlemek ; Meydana gelen suç kalıpları ve serilerinden sorumlu olabilecek şüphelileri tespit etmek ; ve tanımlanmış türlere ve eğilimlere göre gelecekteki suçları öngörmek. İlk hedefle ilgili olarak (kalıpları ve suç serilerini tanımlamak), bir kalıp ancak en az bir suç değişkeninin tekrarı olduğunda var olabilir. Ayrıca, bir kalıp ancak her bir suçun aynı kişi (ler) i işlediğine inandıktan sonra bir seri olarak düşünülebilir. Suç işlemek için kullanılan yöntemler incelenerek bir olayın bir seriye ait olup olmadığı hakkında sonuçlar çıkarılabilir. Taktik suç analizi öncelikle şu anda, günümüzde meydana gelen suçlar ve yakın gelecekte oluşması muhtemel suçlarla ilgilidir. Kolluk kuvvetlerine acil suç sorunlarını belirleme ve anlama konusunda yardımcı olmak için taktik suç analistleri, veri analizlerindeki suçların ne zaman, nerede ve nasıl olduğu ile ilgili bilgilere odaklanır. Özellikle, taktik suç analistleri kalıpları ve suç serilerini tanımlamanın ilk amacını karşılamak için ceza olaylarının yöntemleri hakkında spesifik bilgileri inceler. Örneğin, bir dizi suçu araştıran taktik suç analisti, özellikle binaya veya yapıya giriş noktasına, giriş kazanmak için kullanılan araçlara, suçların meydana geldiği zamana ve çalınan öğelere bakar analistin bir dizi olayı birbirine bağlamasına izin verecek diğer faktörler veya suçun belirli özellikleri ile birlikte.

2.2.2 Stratejik Suç Analizi

Suçun stratejik analizi esas olarak uzun vadeli suç ve düzensizlik sorunları için polis stratejilerinin planlanması ve uygulanmasına odaklanmaktadır. Küresel suç faaliyeti türlerini tanımlar ve zaman içinde nasıl değiştiklerini inceler; suç eğilimleri olarak adlandırılan olayların sıklığını temsil eden eylem yönü, örneğin bir dizi farklı nedenden dolayı değişebilir yılın zamanı. Uzun vadeli sorunlar ve uzun vadeli suçlarda artış veya azalış beklentileri konusunda endişeli. Stratejik analiz aynı zamanda suç, kaynak edinimi ve tahsis çalışmaları üzerine istatistiksel özetlerin hazırlanmasını da içermektedir. Stratejik analiz, "suç eğilimleri" adı verilen uzun vadeli suçlardaki artışları veya azalmaları inceler ; Suç eğilimi, suçun yönüdür ve belirli bir yargı alanındaki veya bölgedeki suç sıklığında herhangi bir değişikliği veya artışı / azalmayı yansıtmaz. Örneğin, vatandaşların araçlarını ısıttığı kış aylarında araba hırsızlığındaki artışı inceleyebilir, böylece çeşitli yerlerde kilitsiz ve gözetimsiz bırakabilirler. Stratejik suç analizi öncelikle toplam verilerin nicel analizinden oluşur. Suç, hizmet çağrısı ve trafik bilgisi gibi cezai ve cezai olmayan bilgilerin aylık, üç aylık ve / veya yıllık derlemeleri toplu olarak analiz edilir. Yani, olayların anlatı açıklamaları gibi nitel veriler yerine tarih, saat, yer ve olay türü gibi genel kategoriler analiz edilir. Irk, sınıf, cinsiyet, gelir, nüfus, yer ve yer türü gibi değişkenler, analiz sürecinde kolluk bilgileri ile birlikte incelenir.

Stratejik suç analizinin iki temel amacı:

- İlaç aktivitesi veya oto hırsızlığı gibi uzun vadeli sorunların tanımlanmasına ve analizine yardımcı olmak.
- İlgili cevapları ve prosedürleri araştırmak veya değerlendirmek için çalışmalar yapmak.

Bu amaçların her ikisi de problem çözme sürecine çok iyi karşılık gelir. Örneğin, stratejik suç analistleri, sıcak havanın sokaktaki insan sayısını artırdığı yaz aylarında agresif panhandler'larla ilgili sorunlara bakabilirler. Başka bir örnek, daha fazla insanın araçlarını çalışır durumda bırakması muhtemel kış aylarında önemli ölçüde artma eğilimi gösteren araba hırsızlığı miktarını azaltmak için bir eylem planı oluşturmayı içerebilir. Stratejik suç analistleri genellikle belirli bir bölgeye veya bir kolluk kuvvetinin yargı yetkisine atanırlar ve bu nedenle suç eğilimlerini, bir bütün olarak örgütten ziyade kendi bölünmeleri içinde tanımlar ve analiz ederler. Bu, suç eğilimlerinin yer değiştirmesi ve suç modellerindeki değişiklikleri açıklayabilecek diğer

itici faktörler gibi şeyleri belirlemek için yetkiler arasındaki çapraz karşılaştırmalara izin verir.

2.2.3 İdari Suç Analizi

İdari suçları analiz etmenin amacı idareye ekonomik, coğrafi veya sosyal bilgi sağlamaktır (Boba, 2005 ; Osborne ve Wernicke, 2003). Bu analiz, bölüm başkanları, komuta ve kontrol personeli ile diğer şehirler, hükümetler, personel ve kamuya açıklayıcı suç bilgilerinin verilmesini içermektedir. Bu raporlar yöneticileri kaynakları belirleyip tahsis ettiklerinde veya vatandaşların toplumdaki suçları ve huzursuzluk sorunlarını daha iyi anlamalarına yardımcı olduklarında destek sağlar. İdari suç analizinin temel amacı, kolluk yöneticilerini ve üst düzey yetkilileri, bölgelerinde meydana gelen suç ve düzensizlik sorunları hakkında ve analitik tekniklerin, suç ve düzensizliğin daha fazla tekrarlanmasını önleme ve önlemeye yönelik analitik tekniklerin nasıl katkıda bulunduğu konusunda bilgilendirmektir. İdari suç analizi genellikle doğrudan kolluk kuvvetleri içindeki üst düzey yetkililere yapılan yazılı raporlar ve sözlü sunumlar şeklindedir. İdari suç analizinin ikincil rolü, halkı, kolluk kuvvetlerinin halihazırda topluluklar içindeki suçu azaltmak için ne kadar çaba sarfettikleri ve çabalarının bugüne kadar ne kadar başarılı olduğu konusunda bilgilendirilmesidir. Suç analistinin bu işlevdeki rolü için çok önemli olan, bilginin ifşa edilmesine ilişkin etik ve gizlilik kurallarına sıkı sıkıya bağlı kalarak kolluk kuvvetlerinin bütünlüğünü korumaktır. Genel olarak, suç analisti belirli proaktif polislik stratejilerini özetleyen bir sunum oluşturmayı isteyecek ve sonuçları belirli bir topluluk ya da bölünmeyi oluşturacak ve daha sonra belirli bir alandan sorumlu üst düzey yetkililerden biri tarafından konuşulacak. Bulgunun hedeflendiği analiz ve izleyicilerin amacı, sunumun son taslağında ortaya çıkanları büyük ölçüde etkilemektedir. Bu nedenle, sunulan bilgiler genellikle tüm analitik ve araştırma sürecinin sadece küçük bir kısmını temsil eder.

2.2.4 Cezai Soruşturma Analizi

Soruşturma suçu, mevcut bilgilerin analizine dayanarak araştırmacılar için şüpheli ve mağdurların profillenmesini içerir. Genel olarak, ne tür bir kişinin belirli bir dizi suç işlediğini varsaymaya odaklanır. Bu tür analizde, suçun niteliğine, davanın gerçeklerine ve mağdurun özelliklerine bağlı olarak bilinmeyen bir suçlunun profillenmesi süreci

olan bir profil oluşturma da vardır. İstihbarat analizinde olduğu gibi, bu analizler öncelikle cinayet ve tecavüz gibi ciddi seri suçları çevreleyen niteliksel verilere odaklanmaktadır. Veriler, temel olarak veya çevreye dahil olan kişiler için bireysel düzeyde toplanır ve analiz edilir. Suç soruşturması analizinin temel amacı, suçlu ve / veya açık davaları yakalamak için olayların içinde ve arasında davranış ve deliller arasında bağlantı kurarak şehir, eyalet ve hatta ulusal sınırları geçen seri suç kalıpları geliştirmektir (Boba, 2005).

Bir suçlunun kişiliği hakkında böyle bir çıkarımda bulunmak ve bilgileri araştırmacıların kolayca erişebileceği ve kullanabileceği bir formatta derlemek cezai soruşturma analistinin görevidir. Şüpheli profiller, suç geçmişi, bilinen ortaklar, bağlantılı adresler ve mülkler ve suçlunun belirli bir şüpheliyle ilgili olabilecek herhangi bir polis güvenliği endişesi ile birlikte uymak zorunda olduğu tüm koşullar dahil olmak üzere belirli bir şüpheliyle ilgili mevcut tüm bilgileri içerir. silah taşıdığı biliniyor). Emniyet teşkilatlarının belirli suçluları hedef alan stratejiler ve taktikler oluşturmalarına yardımcı olurken, aynı zamanda polis tarafından bilinmeyen hedefler için bir başlangıç noktası sağladığından, bu bilgilerin hepsinin el altında olması devam eden bir soruşturma için paha biçilmezdir.

2.2.5 Zeka Analizi

İstihbarat analizi, organize suç, terörizm ve özel soruşturmaların bilgi analizi ve sunumu ile desteklenmesine odaklanmaktadır. Analistler, memurlar için bilginin "işleyicisi" olarak araştırmaları destekleyebilirler. Bir cinayet soruşturmasında, analiz araçları soruşturma bilgilerini düzenlemek ve bunları zaman çizgileri ve ilişkilendirme bağlantı grafikleri şeklinde görüntülemek için kullanılabilir. İlgili bir amaç, analizleri anlaşılması kolay bir çerçeveye yerleştirerek bilgileri birbirine bağlamak, bilgileri önceliklendirmek, ilişkileri belirlemek ve daha fazla araştırma yapılacak alanları tanımlamaktır. Atkin, H. (2011) istihbarat analizi alanında analiz edilen bilgilerin çoğu vatandaşlar tarafından polise bildirilmemekte, kolluk kuvvetleri tarafından toplanmaktadır. İstihbarat analizi geleneksel olarak az çok uyuşturucu ve fuhuş sendikalarını içeren organize suç faaliyetlerine odaklanmıştır. İstihbarat analizinin temel amacı, suç şebekesi üyeleri arasındaki bağlantıları tanımlamak ve nihayetinde bu suçluların yakalanmasına yardımcı olmaktır. Diğer suç analistlerinin aksine, istihbarat analistleri bilinen bir sorunla veya belirlenmiş bir suç ağıyla (yani motosiklet çetelerini

yasaklayın) başlar ve daha sonra önceden belirlenmiş hedefe özgü bilgileri toplamaya, analiz etmeye, derlemeye ve yaymaya devam eder. Bu, suç eğilimlerini ve örüntülerini tanımlamak ve daha sonra bu örüntülerin ve eğilimlerin doğası ve kapsamı hakkında açıklamalar sağlamak için büyük miktarlarda veri aramaktan sorumlu olan diğer suç analistlerinin rollerinin aksine. Suç istihbaratı analizi genellikle çeteler, uyuşturucu kaçakçıları, terörizm, fuhuş halkaları gibi büyük organize suç şebekeleri için veya birçoğu birbiriyle örtüşmeye eğilimli olduğundan (örneğin, birçok çete kuruluşu aktif olarak Uyuşturucu ticareti.

Osborne and Wernicke (2013) bu nedenle, istihbarat analizi çoğunlukla il, ulusal ve uluslararası düzeyde gerçekleşir. İstihbarat analistleri, bilgileri araştırmacılar tarafından kolayca erişilebilecek ve kolayca anlaşılabilecek bir formatta işleyerek ve düzenleyerek polis soruşturmalarını destekler. İstihbarat analistleri tarafından incelenen verilerin çoğunluğu, gözetim, gizli operasyonlar ve telefon dinleme gibi soruşturma teknikleri ve stratejileri aracılığıyla kolluk kuvvetleri tarafından toplanmaktadır. Bu kadar kapsamlı cezai ağları araştırırken, belirli bir şüpheli veya grupta toplanabilecek her türlü bilgi, sadece cezai verilerle değil, bir soruşturmayla da ilgili olabilir. Bu nedenle analistler finansal ve seyahat bilgilerini, aile ve iş ilişkilerini ve analizlerinde telefon kayıtlarından elde edilen verileri içereceklerdir. Tüm farklı veri analistleri türlerini inceleyerek, belirli bir suçun meydana gelmesi için gerçekleşmesi gereken olaylar dizisine ilişkin olarak ilişkileri tanımlayabilir, bilgilere öncelik verebilir, daha fazla araştırma yapılması gereken alanları tanıyabilir ve bağlantılar ve sonuçlar oluşturabilirler. İstihbarat analistleri daha sonra bu bilgileri zaman çizgileri ve ilişkilendirme çizelgeleri şeklinde derler ve görüntüler. Örneğin, yerel bir ailenin evinden küçük bir toplulukta yasadışı uyuşturucu operasyonu yürütülmesi örneğini düşünün. Operasyonun başı, 3 çocuğu tarafından uyuşturucu kaçakçısı olarak çalışan bir dizi kişi ile yardımcı olan ebeveynlerdir. İstihbarat analisti, ilaç evinin adresini verir ve araştırmacı tarafından ikametgahla ilgili insanlar ve araçlar hakkında her türlü bilgiyi bulmasını ister. Analist, ajansın veritabanları aracılığıyla telefon kayıtlarını ve çalışan adları ve adresleri inceleyerek ilgili tüm bilgileri bulabilir ve tüm parçaların nasıl birbirine bağlandığını ve bir suç şebekesi üyelerinin nasıl ilişkilendirildiğini gösteren bir ilişkilendirme çizelgesinde derleyebilir.

2.2.6 İşlem Analizi

Operasyonel analiz, suç azaltma faaliyetinin planlanmasında ve operasyonel hedeflere ulaşmak için kaynakların dağıtılmasında alan komutanlarını ve bölgesel operasyon yöneticilerini destekleyen bir istihbarat ürünü oluşturulmasıdır. Bazı profesyonel kuruluşların bunu farklı bir zeka çalışması seviyesi olarak görmemesi bir hatadır; bir şey varsa, bu suç isti Bu orta seviye operasyon, coğrafi alanlardan sorumlu olan veya icra ekiplerine komuta eden karar vericileri desteklemektedir. Operasyonel analiz, karar vericilerin hangi organize suç gruplarının yaptırıma en savunmasız olduğuna veya bir şehrin hangi bölgelerinin en fazla kaynağa ihtiyaç duyduğuna karar vermesine yardımcı olur. Sınırlı kaynaklara sahip komutanların önümüzdeki birkaç hafta veya ay için ana öncelikleri belirlemesine izin verir ve birkaç tutuklama ile hafifletilemeyen uzun vadeli sorunların büyük bir resmini anlar. hbaratının en hızlı büyüyen alanıdır.

Operasyonel analiz öncelikleri genellikle problem çözme yaklaşımları ve sorun odaklı polislikte kullanılan tip metodolojileri için hedef olarak olgunlaşmaktadır. Ölçek açısından, taktiksel analizin üzerinde durmaktadır, çünkü operasyonel analiz, bir coğrafi alandaki en acil suç azaltma önceliklerini belirlemekle daha çok ilgilidir; ancak karar verenin öncelikleri belirledikten sonra herhangi bir uygulama hedefini desteklemek için taktik analiz kullanılacaktır. İşlem analizi, bir kolluk kuvvetinin kaynaklarını nasıl kullandığını inceler. Dağıtım, hibe fonlarının kullanımı, yeniden dağıtım atamaları ve bütçe sorunları gibi konulara odaklanmaktadır. Birçok ajansta, suç analistlerinden operasyon analizi kategorisine giren bölüm için özel projelere yardımcı olmaları istenmektedir.

2.2.7 Finansal Suç Analizi

Uluslararası olarak kabul edilmiş mali suç tanımı yoktur. Mali suç genellikle bir başkasının mülkiyetinin kendi kişisel kullanımı ve yararına yasa dışı olarak dönüştürülmesini içeren mülke karşı suç olarak tanımlanır. Uluslararası Para Fonu'na göre, mali suç "genel olarak mali kayıpla sonuçlanan şiddet içermeyen herhangi bir suç" anlamına gelir. İngiltere'de, Finansal Hizmet ve Piyasalar Kanunu 2000 (FSMA) mali suç şunları içerir: fra dolandırıcılık veya sahtekârlık içeren herhangi bir suç; bir finansal piyasada yanlış davranış veya bilgilerin kötüye kullanılması veya suç gelirlerinin ele alınması'. Bazı araştırmacılar ve hükümet kurumları mali suçları, mali suistimalleri ve

beyaz yakalı suçları net bir şekilde ayırt etmiyor ve bunları herhangi bir fark olmadan birbirinin yerine kullanıyor. finansal kötüye kullanımın zayıf düzenleyici ve denetleyici çerçeveler ve zayıf vergi sistemleri tarafından teşvik edildiğini ve finansal kötüye kullanımın bir alt kümesi olarak finansal suçun finansal kayıp gerektirdiğini ileri sürmektedir. Beyaz yakalı suçlara örnek olarak kara para aklama, içeriden öğrenenlerin ticareti, dolandırıcılık ve finansal suçla yakalanabilecek piyasa manipülasyonu verilebilir (Jung ve Lee, (2017).

Pickett ve Pickett mali suçu, mali suç, beyaz renk suçu ve sahtekarlık terimlerini birbirinin yerine kullanarak 'normalde güven ihlali içeren yasadışı kazanç için aldatmacanın kullanılması ve faaliyetlerin gerçek doğasının gizlenmesi' olarak tanımlamaktadır. Öte yandan Interpol, mali suçun, Sutherland tarafından ilk kez "mesleği sırasında saygınlık ve yüksek sosyal statüde bir kişi tarafından işlenen" beyaz yakalı suçtan da söz edildiğini belirtiyor Mali suçun, beyaz yakalı suçu içerdiği, birincisinin mesleklerden bağımsız olarak bireyler veya gruplar tarafından işlendiği analiz edilmektedir. Mali suç farklı suçları kapsar: yolsuzluk, kara para aklama, sosyal sahtekarlık ve vergi kaçakçılığı. Bu suçlar ortak bir sosyal mantığı paylaşıyor: toplumlarımızda önemli bir büyüklükte olmalarına ve ekonomik dengeyi ve demokratik değerleri doğrudan tehdit etmelerine rağmen, kayıtsız görünen yüz milyonlarca dolar yönelmiş vatandaş için bir şekilde "görünmez". "kurbansız suç", devletin sanal kurbanı. (Gottschalk (2010)) çok çeşitli mali suçları dört ana kategoriye ayırmıştır (yani Yolsuzluk, sahtekarlık, hırsızlık ve manipülasyon). Bu dört kategorinin neden seçildiğine ve gerekçelendirilmemesine rağmen, yakın tarihli mali suç türleri ayrıntılı bir şekilde gözden geçirildi.

Çizelge 2.2 Finansal suçun Gottschalk tarafından sınıflandırılması

Bozulma	Geri Tepmeler, Rüşvet, Gasp, Zimmete Para Geçirme
Dolandırıcılık	Kimlik, Ipotek, Mesleki
Hırsızlık	Nakit, Entelektüel, Dolandırıcılık
Hile	Aklama, Siber Suç, Teklif Arma, Içeriden Öğrenenlerin Ticareti

Yolsuzluk, kamu veya özel bir fonksiyon tarafından kolaylaştırılan kişisel yararları kabul etme veya sunma eylemidir. Kara para aklama, vergi kaçakçılarından örgütlü uyuşturucu kaçakçılığı gruplarına kadar değişen mali suçların en büyük suçudur ve suç

fonlarından cezasızlıkla kâr etmeyi amaçlamaktadır. Sosyal sahtekarlık, sosyal damping ve kayıt dışı çalışma ve hileli sosyal güvenlik kapsamına (ONSS'nin yanlış beyanları, yanlış maaş bordrosu, banka kredilerine erişim, aile birleşimi vb.) odaklanmaktadır. Sonuç olarak, bazen insan kaçakçılığı ve kaçakçılığı ile bağlantılıdır, ancak özellikle işçilerin ve serbest çalışanların haklarına zarar verir ve haksız rekabete yol açar.

Kara para aklama: Kara para aklama, kriminolojide büyük bir suç olarak kabul edilir ve toplumdaki en büyük grup suçlarından biri olarak tanımlanır. Küresel ekonominin gelişmesi, internetin artan uygulamalarının ve e-ticaretin ilerlemesiyle, kara para aklama suçlarının (MILC'ler) daha yaygın, araştırılması daha zor ve sağlıklı gelişimine daha zararlı olacağı tahmin edilmektedir. ekonomi ve finansal sistemlerin istikrarı.

FATF bunu "suçluların yasadışı kökenlerini gizlemek için işlemek" olarak tanımlarken, ABD Hazine Bakanlığı bunu "terör örgütleri de dahil olmak üzere suçluların yasadışı kişilerin gelirlerini, kaynaklarını veya mahiyetini gizlemeye çalıştıkları finansal işlemler olarak daha ayrıntılı olarak açıklamaktadır". Kara para aklama takip edilmeden önce kesin bir suç bulunmalıdır. Dolayısıyla, kara para aklama, paranın suç gelirleri olarak kabul edilmemesi halinde kabul edilmez. Söz konusu suç fiziksel suçlar olabilir, ancak IMF (2001) genellikle finansal suçlar olduğunu ileri sürmektedir. Yüklü suç türleri, bir ülkenin bağlamına ve yargı yetkisine göre değişir. Genellikle kredi kartı sahtekarlığı, sahtekarlık veya ipotek sahtekarlığı ve kimlik avı gibi diğer suçları kolaylaştırmaya kararlıdır. Genellikle mali suçların çoğunun kolluk kuvvetleri tarafından ifşa edilmemesi için kara para aklamaya neden olduğu ileri sürülmektedir. Vergi kaçakçılığı, gelir vergisi veya KDV yükümlülüklerine uyulmamasına neden olur.

2.2.8 Kimlik hırsızlığı

Johnsen, (2016) çalışmasdında kimlik hırsızlığından gelen bilgilere dayanarak, dolandırıcıların bir hikaye oluşturması veya potansiyel kurbanları dolandırmak için birini taklit etmesi Gottschalk (2010) kimlik hırsızlığının kimlik sahtekarlığını mümkün kıldığını iddia etmektedir. Kimlik hırsızlığı bir izinsiz giriştir veya daha fazla bilgi toplanması neredeyse tespit edilemeyen kimlik hırsızlığına da izin verir. Kimlik hırsızlığı saldırıları, bir kişinin veya kuruluşun dijital kimliğiyle ilgili gizli ve kişisel bilgilerin, ikincisinin onayı olmadan, yetkili bir kullanıcı olarak poz vermek için kullanılmasını içerir.

Böyle bir saldırının tipik örneği IP adresi sahteciliğidir. İnternetin sanal ortamında, güvenli erişim rozeti ile aynı karaktere ve aynı kimlik doğrulama gücüne sahiptir ve bilgisayar korsanının bilgisayar hedefinin kontrolünü ele geçirmek için bir ihlal açmasına izin verir. Bu nedenle, bu tür saldırılarda hem bilgisayar saldırılarını hem de kimlik hırsızlığı katılımcılarını takip etmek akla yatkın görünmektedir. Elektronik ticaretin gelişmesiyle birlikte, kimlik hırsızlığı sıradan ve çok karlı, çok zarar verici hale geldi. Bu fenomenle karşı karşıya kalan Devlet, aynı zamanda şirketler ve bireyler de tepki göstermelidir. Bilgisayarın izinsiz giriş riskini azaltmak için yukarı doğru bilgisayar koruması kullanmak önemlidir, ancak aynı zamanda bilgisayara izinsiz giriş eylemlerini gerçeğe uygun değerine dayandıracak yasal bir aşağı akış kapsamı da içerir. Yasal koruma sistemi : Yasama organı, Ceza Kanunu ile bilgisayar sahtekarlığına özgü yeni suçlamalar başlatmıştır. Saldırının korunup korunmadığına göre ayırt eder. Verilen cezalar, izinsiz girişin bilgisayarda bir etkisi olup olmadığına bağlı olarak değişir. Söz konusu sistem, yani basit veya hasarsız bir saldırı ve hasarlı bir saldırı.

Kişisel verilerin korunması: Elektronik araçlar önemli ölçüde gelişti. Bilgisayarlar, cep telefonları yaygınlaştı ve İnternet ve sosyal ağlar iletişim tarzlarında devrim yarattı. Gerçekten de, elektronik ticaretin yükselişiyle birlikte veriler önemli bir ekonomik konu haline gelmiştir. Gerçekten de, ticari bir zorunluluk olan bu dijital ortamda, bu "hammaddeler" agresif pazarlamanın yeni silahları haline gelir. Toplandıktan sonra, şirketlerin piyasaları daha iyi tanımasına, gerçek veya potansiyel müşterilere her türlü hedefleme veya av uygulamalarına izin verir. Tehlikelerin bu evrimi karşısında, bu kişisel veriler bugün korunmaktadır.

2.2.9 Siber Suç Analizi

Hosseinkhani ve Naniz (2014) çalışmasında siber suçlar, yasadışı faaliyetlerin bilgisayar ve internet kullanılarak gerçekleştirildiği anlamına gelir. Siber suçlar temel olarak iki ana kategoriye ayrılabilir. Bunlardan biri ağı, izinsiz giriş, ağ sistemini tahrip etme, vb. Şiddet, pornografi, dolandırıcılık bilgileri internette her yerde görülebilir. Bu bağlamda, Web madenciliği, internette patlayıcı ve yapılandırılmamış bilgi toplama ve analiz için yeni bir araştırma yönüdür. Cezai web verileri her zaman hukuk yönetimi için değerli ve uygun bilgiler sunar. Geniş çaplı suç web verilerinin farklı kapasitelerinin değerlendirilmesi her zaman çok zordur, bu nedenle hukuk yönetimi için en dikkate değer görevlerden biridir. Suçlar, Web madenciliğinin bir çözüm olarak

geldiği verilerden yararlı bilgiler elde etmek için gelişmiş analitik yöntemlerin gerekli olduğu cinayet ve tecavüz kadar aşırı olabilir. Kesinlikle, bir suç fenomeniyle karşılaşan etkili faktörlerden biri, insanların sosyal yaşam koşullarıdır, bu nedenle suç analizi bilgisine etkili bir mücadele aracı olarak ihtiyaç duyulmaktadır. Suçluluk ile ilgili veriler ve suç karmaşıklığı ve bunlar arasındaki maddi olmayan ilişkilerin varlığı nedeniyle, veri madenciliği kriminologlar arasında hızla büyüyen bir alanda yapılmaktadır. Gibi suçlu nesneler olarak kabul edenlerdir. İnternet, siber suçlar için verimli bir zemin yarattı. Tayvan bölgesi ve Japonya'dan araştırmacılar tarafından yürütülen bir istatistik raporuna göre, yasadışı ağ kullanım vakalarının sırayla en yaygın oranı: internet pornografisi, internet sahtekarlığı, yasadışı mal ticareti, yıldırma ve gasp, yasadışı izinsiz giriş, hakaret ve iftira. Birçok gerçek, internetteki bilgilerin sadece geleneksel idari modeller aracılığıyla yönetilmesinin yeterli olmadığını kanıtlamıştır. Web madenciliği araştırmalarının odak noktası yeni web madenciliği teknikleri geliştirmek ve bunları temsil edecek metinlerin özelliklerini çıkarmaktır. Emniyet müdürlüklerinde, suçla ilgili büyük miktarda veri bulunmaktadır.

Ayrıca, suç olaylarını keşfetmek, tanımlamak ve tahmin etmek için sistematik bir yaklaşımın kullanılmasını içerir ve girdisi, suç değişkenlerinde atanmış bilgi ve veriler içerir ve çıktı, bilgi çıkarma, analitik ve soruşturma sorularının cevabını ve sonuçların görselleştirilmesini içerir. Gözaltına alınan vatandaşlar, tehlike altındaki çocuklar, işletmeler mahvoldu, devletler tehdit etti, siber suçlular İnternet büyüdükçe sahiplerini uzattı. Onları görmüyoruz, onları tanımıyoruz, onlara güvenmiyoruz ve bu onların gücüdür. Yine de hepimiz endişeliyiz. İster fikir, casusluk, kimlik hırsızlığı, terörizm, taciz, suçluluk, finansal sahtekarlık veya çeşitli suçluluk biçimlerinin manipülasyonu olsun, siber suç toplumu bütünüyle. İnternet tarafından sunulan çeşitli hizmetleri kullanarak, her İnternet kullanıcısı kendisini suç kaynaklı bir tehdide maruz bırakabilir ve bir suçun hedefi veya aktörü, çoğu zaman istemsiz bir suç haline getirebilir. Haberler birkaç yıldır düzenli olarak gösterdiği için siber suçlar artık bir gerçek. Bireyler, kuruluşlar ve devletler için az ya da çok önemli sonuçları vardır. Genel anlamda, siber suç, genellikle bir ağa bağlı bir bilgisayar sisteminin kullanımı ile işlenebilen veya kolaylaştırılabilen suç grubudur. Böylece tanımlanan, çok farklı suçları ifade edebilir.

Johnsen, J. W. (2016) istatistiksel çalışmaları aşağıdakileri ayırt etmeyi gerektirir: Bilişim sistemleri ve otomatik veri işleme sistemleri (hileli erişim, bir sistemin değiştirilmesi, servis reddi saldırısı ...) ile ilgili, bilgisayar ağlarının, özellikle de

internetin gelişmesinden kaynaklanan suçlar. Yeni bilgi ve iletişim teknolojileriyle (NICT'ler) evrimleşmiş olabilecek veya onlar tarafından kolaylaştırılabilecek ve böylece yeni bir suç vektörü teşkil edebilecek "geleneksel" suç biçimlerine bağlı suçlar. Bu ikinci kategori, örneğin, dolandırıcılık, yeni İnternet'te (çevrimiçi kredi kartlarının hileli kullanımı, kimlik avı, vb.) ortaya çıkmış formları, şimdi yenileriyle yayılabilecek her türlü tehdit ve hakaret. Elektronik iletişim (mesajlaşma, forumlar, sosyal ağlar ...), ağlar ve internet tarafından da kolaylaştırılan çocuk pornografisi görüntülerinin yayılması, vb.

Birkaç yıldır Amerika Birleşik Devletleri ve Kanada siber güvenliği ulusal bir öncelik haline getirmiştir. Bu saldırılar tarafından bilgisayar saldırılarına karşı savaşmak için önemli aygıtlar uygulanmaktadır. Bununla birlikte, siber suç sınırların üstesinden gelme özelliğine sahiptir. Dolayısıyla bütün devletler siber suçlarla ilgilenmektedir. Siber suçların yol açtığı zararın ciddiyeti, bunlar arasındaki küresel işbirliğini haklı çıkarmaktadır. Budapeşte Konvansiyonu'nun kabul edilmesi ve bu belgenin Avrupa dışı Devletler tarafından çıkarılması, siber suçlarla mücadele için işbirliğinin artırılması gerektiğine tanıklık ediyor.

Bazı çok taraflı örgütler de, tartışmaları siber savunmaya daha çok yönelmiş olan NATO başta olmak üzere bilgi sistemlerinin güvenliğini vurgulamaktadır. Çin ve Rusya bilgi sistemlerinin düzenlenmesini sıkı bir şekilde savunmaktadır. Gerçekten de, 2009 yılında, Birleşmiş Milletler'de bilgi güvenliği için bir davranış ilkesi benimsemesi yönünde bir karar önerdi. Bu girişim başarılı olamadı. Bu anlamda, bu karar, düzenlemenin çözümü olarak değil, bir ifade ve ifade özgürlüğünün sınırlandırılması olarak algılanmıştır. Hükümet ve politikacılar tek bir sesle siber suçların belasını kınıyor ve tüm devletler arasında uyumlu eylemler öneriyorlar. Bununla birlikte, daha önce kabul edilen önlemler ve girişimler çok seyrek ve gerçek işbirliğini ve eylemlerin uyumlaştırılmasını pekiştirmeye izin vermiyor. Bu zorunluluğa rağmen, diyaloglarda zorluklar ortaya çıkıyor. Bugün büyük güçler ortak bir politika oluşturmak için gereken fikir birliğini bulamıyorlar. Fikir birliği eksikliği, uluslararası girişimleri yönetmektedir. 1980'lerden beri çeşitli girişimler uygulanmaktadır. Kararlar böylece Birleşmiş Milletler ve Uluslararası Telekomünikasyon Birliği düzeyinde kabul edilir. OECD veya NATO gibi diğer kuruluşlar da internetteki faaliyetlerin düzenlenmesini başlatmış ve siber suçları azaltmaya çalışmıştır.

Suç analizi süreci: Analiz sürecinde, bir sorunun en iyi şekilde önce kapsamını inceleyerek ve ardından bileşenlerini tanımlayarak çözüldüğüne dair temel bir varsayım vardır. Başka bir varsayım ve suç analizine özgü bir varsayım, suçluların genellikle eylemlerini tekrarlayan bir tarzda ve tarzda gerçekleştireceğidir. Bu nedenle, analiz süreci bir suçun işlenme biçiminin tanımlanmasını, suçun benzer suçlarla karşılaştırılmasını ve suçun bilinen suçluların MO "stili" ile karşılaştırılmasını içermelidir. Suç faaliyetine uygulanan analiz süreci, birbirine bağlı beş fonksiyonun adım adım özel bir dizisidir:

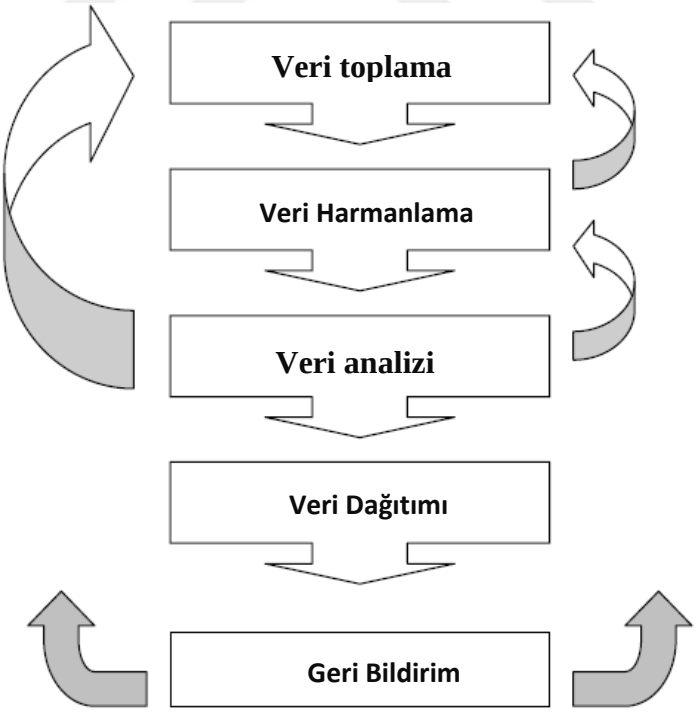
• Suç verilerinin toplanması

• Suç verilerinin harmanlanması

• Analiz

• Analiz raporlarının yaygınlaştırılması

• Geri bildirim ve değerlendirme



Şekil 2.1 Suç analizi süreci

Veri toplama: Bilgi ve veri toplama, suç analizi sürecinde en önemli adımdır; veri olmadan süreç var olamaz. Suç analistlerine sunulan veriler ve bilgiler birçok farklı biçimde gelir, ancak iki genel kategoriye ayrılabilir: iç bilgi kaynakları ve dış bilgi kaynakları. Dahili bilgi kaynakları, suç raporları, tutuklama raporları, sokak kontrol raporları, saha görüşme raporları, rezervasyon raporları, sevk kayıtları, mağdur ve tanık ifadeleri ve kanıt raporları gibi şeyleri içerir. Harici bilgi kaynakları, diğer kolluk kuvvetleri, motorlu taşıtlar bölümü, şartlı tahliye / denetimli serbestlik daireleri ve diğer federal kurumları içerir. Toplama aşaması da etkili olabilmesi için planlanmalı ve odaklanmalıdır. Yani, suç verileri yalnızca tekrarlanan sıklıkları veya kalıpları nedeniyle analiz edilebilecek suçlar üzerinde toplanmalıdır. Örneğin cinayet, aynı suçlu tarafından nadiren tekrarlanan bir suçtur. Bu nedenle, bu tür suç verilerini sürekli olarak toplamak genel bir amaca hizmet etmeyecektir.

Veri toplama, ilgili değişkenler hakkında bilgi toplama ve ölçme, belirtilen araştırma sorularını cevaplamak, varsayımları test etmek ve sonuçları değerlendirmek için kurulmuş sistematik bir yöntemdir. Veri toplama adımında, haber siteleri, bloglar, sosyal ağ, RSS beslemeleri vb. Suç verileri yapılandırılmamış bir veridir, çünkü belgenin alan, içerik ve boyutu bir belgeden diğerine farklılık gösterebilir, daha iyi bir seçenek şemadan daha az veritabanına sahip olmaktır. Gibi farklı web sitelerinden veri topluyoruz. Toplanan veriler daha sonraki işlemler için veritabanında saklanıyor.

Veri Harmanlama: Veri harmanlama, suç analizi sürecinde bir sonraki adımdır. Ham veriler kendi başına nadiren çok değerlidir. İlk dosyalar oluşturulduktan sonra, analist bunlara veri ekleyecektir. Benzer öğeler birlikte değerlendirilinceye kadar analist verilere anlam sağlayamaz. Veri toplama ve analiz bu amacı gerçekleştirir.

Dosya Geliştirme Süreci : Veri çıkarma, her suç raporundan benzer verilerin alındığı ve ayrı ve farklı dosyalar (yani Şüpheli dosyalar, araç dosyaları vb.) Geliştirmek için kullanıldığı bir işlemdir. Kolluk belgeleri birçok veri türü içerir. Suç analizi için en yararlı olduğu düşünülen kategoriler aşağıdakilerle ilgilidir:

1. Coğrafi faktörler
2. Zaman faktörleri
3. Kurban tanımlayıcıları
4. Mülkiyet kaybı tanımlayıcıları
5. Fiziksel kanıt tanımlayıcıları

6. Spesifik MO faktörü
7. Şüpheli tanımlayıcılar
8. Şüpheli araç tanımlayıcıları

Veri Analizi Türü: Veri analizi, tanınabilir suç faaliyeti örüntülerinin geliştirilmesi ve suçluların tanımlanması ile sonuçlanan bilgilerin incelenmesi ve işlenmesi olarak tanımlanabilir. İki tiptir:

1. Modus Operandi (MO) Patern Tespiti ve Korelasyonu
2. İstatistiksel analiz

Suç örneği türleri: Verileri analiz etme ve bunlardan sonuçlar çıkarma görevi, bir dekupaj testeresi bulmacasını bir araya getirmek gibidir. Analist, çok çeşitli kaynaklardan toplanan birçok veri elemanı ile çalışmaktadır. Amaç, gerçek bir suç faaliyetinin resminin belirlenebilmesi için verileri birleştirmektir. Resimler kendilerini suç kalıpları şeklinde gösterir. Aşağıda açıklandığı gibi, bunlar iki türdür:

Coğrafi konsantrasyon kalıpları, aşağıdakiler temelinde tanımlanan kalıpları ifade eder:

- Suç türünün benzerliği (ör. Ticari hırsızlık).
- İyi tanımlanmış coğrafi alanlarda çoklu oluşumlar.

Spesifik ve tekrarlayan MO paternleri, aşağıdakiler temelinde tanımlanan paternleri ifade eder:

- Şüpheli ve / veya şüpheli araç tanımlarının benzerliği
- Eşsiz MO özellikleri.

Suç Düzeni/Serisi Tespiti - Manuel Teknikler: Analist, suç olaylarının meydana geldiği yerleri belirtmek için haritaya renk kodlu noktalar yerleştirirse coğrafi konsantrasyon modelleri tanımlanabilir. Nokta haritaları olarak bilinen cihazlar, suç modellerinin görsel olarak tanımlanmasını kolaylaştırır.

Veri Dağıtımı: Analiz aşaması tamamlandığında, analist sonuçlarını kolluk kuvvetleri, polis yönetim ekipleri, vatandaşlar, öğrenciler, diğer analistler ve medya da dahil olmak üzere bir dizi farklı kitleye dağıtmalı ve iletmelidir. Bu nedenle, analistin sonuçlarını yayılmak için hazırlarken bir dizi önemli faktörü dikkate alması önemlidir. Her şeyden önce, verilerin sunumu hedeflendiği kitlenin bilgi düzeyine göre ayarlanmalıdır. Ayrıca

dağıtılan veya sunulan sonucun sadece konuyu veya konuyu elinizde taşımak için gerekli bilgileri içermesi de önemlidir. Analistler tarafından yürütülen çalışmaların çoğu perde arkasında (yani veri harmanlama) yapılır ve tüm sürecin analitik sonuçların yaygınlaştırılmasına ve sunumuna dahil edilmesi gerekli değildir. Daha önce belirtildiği gibi, bilgi paylaşımı, suç analizi disiplininin içsel bir bileşenidir ve bilginin paylaşılmasını ve daha sonra disiplinin geliştirilmesini kolaylaştırmak için yaygınlaştırma süreci esastır. Suç analizi sürecinin son aşaması, suç analizinin iyileştirilmesi ve ilerletilmesi için çok önemlidir. Analizlerin kalitesi, kullanılan verilerin niteliği veya analitik ürünlerin karar vermede yararlılığı hakkında geri bildirim sağlanabilir. Örneğin, genel kamuoyu gibi kolluk kuvvetleri dışında bir izleyici, büyük olasılıkla belirli bir suç türü ile polisle ilgili diğer terminoloji ve kısaltmalar arasındaki farklılıklar hakkında bir açıklama ve açıklama gerektirecektir. Tersine, kolluk kuvvetleri ve adalet bakanlığı çalışanlarından oluşan izleyici bu tür açıklama ve açıklamalara ihtiyaç duymayacak ve işlerinin nasıl düzgün bir şekilde yapılacağına dair bir ders gibi görünebilir. Suç analistleri, ceza infaz kurumlarına cezai soruşturmalarda yardımcı olmanın yanı sıra, ürünlerinin diğer analistlere, araştırmacılara ve polis uygulayıcılarına dağıtılmasıyla suç analizi disiplinindeki bilgi birikiminin artmasına da katkıda bulunmaktadır. Veri dağıtımı, yönlendirilmiş devriye ve taktik eylem planları geliştirmek, devam eden soruşturmalar ve suç önleme çabalarına yardımcı olmak ve genel idari, operasyonel ve organizasyonel planlama görevlerinin yerine getirilmesini kolaylaştırmak için kullanılan suç analizi bilgilerinin dağıtımı olarak tanımlanabilir.

Veri Yaygınlaştırma Hedefleri**:** Veri dağıtımı ile ilgili iki hedef vardır. Bunlar:

- Başkaları tarafından kullanılabilecek veya üzerinde işlem yapılabilecek bilgiler sağlamak
- Organizasyon içindeki ve dışındaki insanlar arasındaki iletişimi geliştirmek.

Ürün Geliştirme Planlama Konuları: En sık, veriler analist tarafından geliştirilen yazılı ürünler aracılığıyla yayılır. Bazıları bölüm içindeki belirli kişi veya birimlerin yararına yayınlanırken, diğerleri kuruluşun tüm üyeleri tarafından kullanılabilir. Yeni suç analistleri ve amirleri için ürün geliştirme sürecini acele etmek nadir değildir. Bununla birlikte, bir şeyleri çıkarmak için acele ettiklerinde, genellikle amaçlanan müşterilerin ihtiyaçlarını gerçekten karşılayan bilgiler veremezler. Bu, aşağıdaki soruların cevaplarını alarak önlenebilir:

Bilgiyi hedefleyen kullanıcılar kimlerdir? Diğerlerinin yanı sıra şunları içerebilir:

- Devriye memurları ve devriye amirleri
- Müfettişler ve soruşturma amirleri
- Taktik aksiyon takımları
- Eğitim memurları
- İletişim / sevk personeli
- Bütçe/diğer idari analistler
- Yöneticiler ve diğer komuta personeli
- Zalimler, denetimli serbestlik ve şartlı tahliye memurları ve diğer kurum dışı personel.

Bilgilendirme ihtiyaçları nelerdir ? Diğerleri arasında aşağıdakilere ihtiyaç olabilir :

- Genel Suç Özeti Bilgisi
- Suç Düzeni Bilgisi
- Suç Serisi Bilgileri
- Suç Eğilimi Bilgileri
- Süpheli Aranıyor
- Çalıntı Araç Bilgisi
- Garanti Belgesi
- Piyon Bilgisi

Hangi ürünler bu ihtiyaçları karşılayacak? Diğerlerinin yanı sıra, dikkate alınacak ürünler şunlar olabilir :

- Günlük Bültenler
- Suç modeli / suç serisi bültenleri
- Suçlu bültenlerini bilir
- Haftalık istatistiki suç özetleri

Geri Bildirim : Geri bildirim, suç analizi ürün ve hizmetlerinin doğruluğunu, güvenilirliğini, geçerliliğini, güncelliğini ve genel kullanışlılığını belirlemek için uygulanan resmi ve gayri resmi iletişim süreçlerinin oluşturulmasından kaynaklanan bilgiler olarak tanımlanabilir. Veri değiştirme alt döngüsü gibi, analitik ürünler kullanan kişilerden gelen geri bildirimler, bir bütün olarak süreçte daha fazla değişikliğe ve

iyileştirmeye yol açabilir. Analitik ürün kullanıcıları ile doğrudan iletişimden elde edilen bu bilgilere dayanarak, analist çalışmalarını sürekli iyileştirmek için prosedürlerini değiştirebilir ve yöntemlerini iyileştirebilir.

Geribildirim ihtiyacı: Bir suç analizi programının oluşturulması ve sürdürülmesinde önemli miktarda kaynak harcanmaktadır. Bu nedenle, suç analizinin de değerlendirilmesi gereken bir nokta vardır.

Resmi Geribildirim Yöntemleri: Resmi yöntemler Bilgi Talebi ve Geri Bildirim formlarının geliştirilmesini içerir. Suç analiz biriminden bilgi veya hizmet talep edildiğinde veya sağlandığında bunlar tamamlanmalıdır. Geri bildirim süreci yöneticilerin ve birim üyelerinin bu görevi yerine getirmelerine yardımcı olur.

Gayri Resmi Geribildirim Yöntemleri: Gayri resmi geri bildirim yöntemleri, ünitenin ve sağladığı ürünlerin etkinliğini de değerlendirir. Basit bir yöntem çöp kovası anketidir. Bu, sadece malzemenizin ne kadarının atıldığını görmek için bölüm boyunca çöp kutularına bakmayı içerir. Bu temanın bir değişkeni, memurlar ve dedektifler alan için ayrıldıktan veya masalarına döndükten sonra devriye ve soruşturma brifing alanlarında kaç ürün kaldığını not etmektir. Malzemenin çoğu atılırsa veya geride bırakılırsa, içerdiği bilgilerin başkaları için çok az değerli olduğunun bir göstergesi olabilir. Birincisi, suçluların yakalanması, suç kalıplarının durdurulması vb. Suç analizi bilgilerinin kullanımı ile doğrudan bağlantılı olamaz. Diğer değişkenler bu sonuçlardan sorumlu olabilir veya bu sonuçlara katkıda bulunabilir. İkincisi, birim bilgi sağlamasına rağmen, birim bu bilgilerin kullanımı üzerinde çok az kontrole sahiptir. Dolayısıyla yukarıdaki faktörlerin değerlendirme kriteri olarak kullanılması, birimin toplam değerinin belirlenmesinde çok az yardımcı olabilir.

Değerlendirme: Değerlendirme, bir programın performansını değerlendirmek, değerini belirlemek ve sürekliliğini veya ortadan kaldırılmasını haklı çıkarmak için uygulanan bir süreçtir. Bir suç analizi programı, kurumun yöneticileri için belirlenen amaç ve hedeflere ulaşması temelinde değerlendirilmelidir. Suç analiz birimlerinin klerens oranları, suç oranları, tutuklama oranları vb. Faktörlere göre değerlendirilmesinden iki nedenden dolayı kaçınılmalıdır.

2.3 Kolluk kuvvetlerinde bir suç analistinin rolü

Austin ve Martensen, (1973) bir emniyet teşkilatında merkezi odak, hem polise bildirilen hem de bildirilmeyen suçtur. Dolayısıyla, analiz edilen merkezi veri türü suçtur ve tutuklamalar, suçlular, mağdurlar, mülk ve kanıt gibi onu çevreleyen bilgilerdir. Suçun yanı sıra, kolluk kuvvetleri diğer birçok konuyu ele alır ve böylece diğer birçok veri türünü toplar. (LISC) ve United States of America. (2017) suç analistleri için genellikle mevcut olan kolluk kuvvetlerine örnek olarak hizmet çağrıları (ör. Gürültü şikayetleri, hırsızlık alarmları, şüpheli faaliyetler), trafik bilgileri (örn. Kazalar ve alıntılar), vatandaşların algıları (örn. Suç korkusu, suç) önleme davranışı, polis memnuniyeti), mağduriyet, denetimli serbestlik kayıtları ve şartlı tahliye bilgileri.

Goldberg, Champagne ve Singleton (2007) suç analistlerinin rolleri ve sorumluluklarının ajanstan ajansa değiştiği ve standart bir suç analizi uygulamasının henüz oluşturulmadığı göz önüne alındığında, bir suç analistinin günlük olarak ne yaptığını doğru bir şekilde açıklamak güçtür.

Çizelge 2.3 Öngörücü Teknolojilerin Kolluk Kuvvetleri Kullanımı

Sorun	**Konvansiyonel Suç Analizi (düşük ila orta veri talebi ve karmaşıklığı)**	**Öngörücü Analiz (büyük veri talebi ve yüksek karmaşıklık)**
Tarihsel suç verilerini kullanma	Suç haritalama (etkin nokta Kimlik)	Gelişmiş etkin nokta tanımlama modelleri ; riskli arazi analizi
Bir dizi ek kullanma veri (ör. 911 çağrı kaydı, ekonomi)	Bir elektronik Çizelge programında oluşturulan temel regresyon modelleri	Regresyon, sınıflandırma ve kümeleme modelleri
Artmış risk muhasebesi yeni bir suçtan	Son suçları derhal çevreleyen alanlarda artan risk varsayımı	Neredeyse tekrarlanan modelleme
Suç riski en fazla olan bölgelerin ne zaman olacağını belirleyin	Belirli bir alandaki suçların sıklığını zamana / tarihe (veya belirli olaylara) göre grafik / haritalama	Konumsal analiz yöntemleri
Suç riskini artıran coğrafi özellikleri belirleyin	Suç olaylarının en sık görüldüğü yerleri bulma ve çıkarım yapma	Riskli arazi analizi

Çizelge 2.4 Tahmin Etme

Sorun	**Konvansiyonel Suç Analizi**	**Öngörücü Analitik**

	(düşük ila orta veri talebi ve karmaşıklığı)	**(büyük veri talebi ve yüksek karmaşıklık)**
Suç grupları arasında şiddetli bir salgın riski yüksek bulun	Gelen çete / suç istihbarat raporlarının manuel olarak gözden geçirilmesi	Neredeyse tekrarlanan modelleme (son gruplar arası şiddet konusunda)
Suçlu olabilecek kişileri belirleyin : Denetimli serbestlik ve şartlı tahliye edilenler en büyük yeniden yerleştirme riski altındadır Yaralanma veya ölüm riski yüksek olan aile içi şiddet vakaları Gelecekteki suç davranışı veya şiddet riski en yüksek olan ruh sağlığı hastaları	Bilinen risk faktörlerini özetleyen klinik araçlar	Risk faktörlerini kullanan regresyon ve sınıflandırma modelleri

Çizelge 2.5 Yasa Uygulamalarının Öngörücü Teknolojilerin Kullanımı: Fail Kimliklerini Tahmin Etme

Sorun	**Konvansiyonel Suç Analizi (düşük ila orta veri talebi ve karmaşıklığı)**	**Öngörücü Analiz (büyük veri talebi ve yüksek karmaşıklık)**
Şüphelileri, mağdurun ceza geçmişi veya diğer kısmi verilerini (ör. Plaka numarası) kullanarak tanımlayın	Suç istihbarat raporlarını elle inceleme ve çıkarımlar çizme	Bilgisayar destekli sorgular ve istihbarat ve diğer veritabanlarının analizi
Hangi suçların bir serinin parçası olduğunu belirleyin (büyük olasılıkla aynı fail tarafından işlenir)	Suç bağlama (bir seride olduğu bilinen suçların niteliklerini diğer suçlarla karşılaştırmak için bir Çizelge kullanın)	Suç bağlantısı yapmak için istatistiksel modelleme
Bir failin büyük olasılıkla bağlantı noktasını bulun	Bir dizi suçun yakınında ve arasında alanları bulma	Coğrafi profil oluşturma araçları (en olası noktaları istatistiksel olarak çıkarmak için)
Olay yeri etrafındaki sensör bilgilerini	Manuel istekler ve sensör verilerinin gözden geçirilmesi	Bilgisayar destekli sorgular ve sensör

kullanarak şüphelileri bulun (GPS izleme, plaka okuyucu)		veritabanlarının analizi

Çizelge 2.6 Yasa Uygulamalarının Öngörücü Teknolojilerin Kullanımı : Suç Mağdurlarının Öngörülmesi

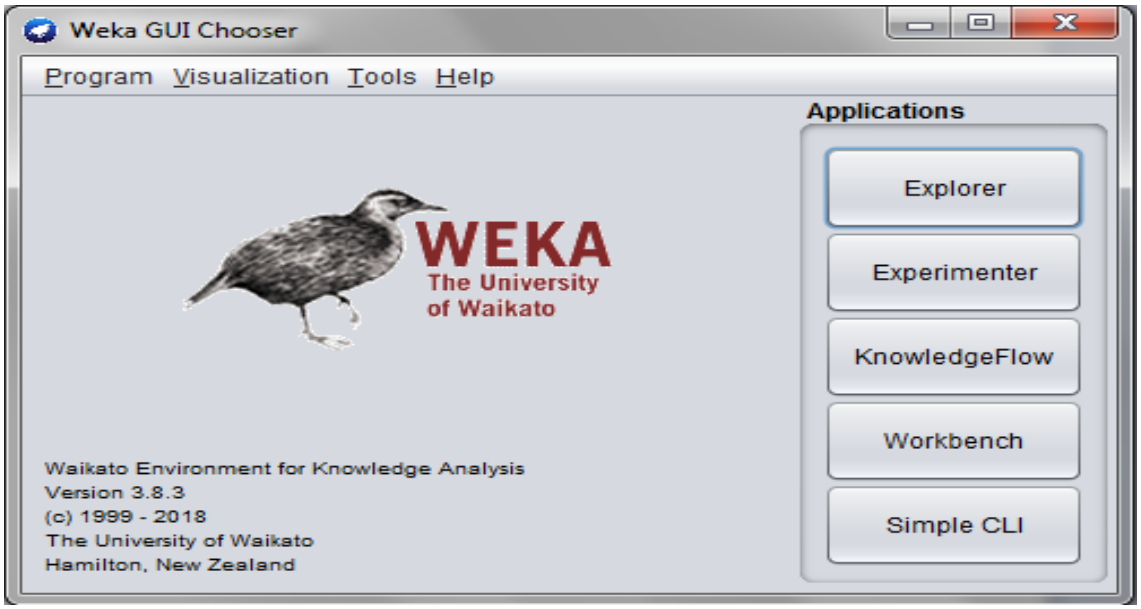
Sorun	**Konvansiyonel Suç Analizi (düşük ila orta veri talebi ve karmaşıklığı)**	**Öngörücü Analitik (büyük veri talebi ve yüksek karmaşıklık)**
Çeşitli suç türlerinin mağduru olabilecek grupları belirleyin (savunmasız nüfuslar)	Suç haritalama (suç tipi etkin noktaları belirleme)	Suç türlerini etkin noktaya göre tanımlamak için gelişmiş modeller ; riskli arazi analizi
Risk altındaki yerlerden doğrudan etkilenen kişileri belirleyin	En sık rastlanan suç sitelerini manuel olarak grafikleme veya haritalama ve bu konumlarda olma olasılığı en yüksek olan kişileri belirleme	Suç mahalleri oluşturmak ve işçileri, sakinleri ve bu mahallleri sıklaştıran diğer kişileri belirlemek için gelişmiş suç haritalama araçları
Mağduriyet riski altındaki kişileri belirleyin (örneğin, yüksek riskli suç davranışıyla uğraşan kişiler)	Tekrarlanan suç faaliyetinde bulunduğu bilinen kişilerin ceza kayıtlarının gözden geçirilmesi	Risk altındaki tekrarlayan suçluları tanımlamak için yerel ve diğer erişilebilir suç veritabanlarında kullanılan gelişmiş veri madenciliği teknikleri

3. MATERYAL VE METODOLOJİ

3.1 Materyal Weka

Kalmegh. (2015) çalışmasında weka Yeni Zelanda'da University of Waikato'de geliştirildi. Bilgi Analizi sistem için Waikato Çevre İçin isim standı Java ile yazılmış ve GNU Genel Kamu Lisansı şartları altında dağıtılmaktadır. Neredeyse her platformu üzerinde çalışan ve Linux, Pencere ve Macintosh işletim sistemleri altında ve hatta kişisel dijital asistan üzerinde test edilmiştir. Bu, giriş verilerinin hazırlanması istatistiksel öğrenme şemaları değerlendirilmesi ve giriş verileri ve öğrenme sonucunun görselleştirilmesi de dahil olmak üzere maden deneysel verilerin tüm süreç için kapsamlı destek sağlamaktadır. Tüm algoritmalar Arff biçiminde tek ilişkisel Çizelge şeklinde onların girişini alır. Bu öncesi ve sonrası işleme yöntemleri ile birlikte ve herhangi bir veri kümesi üzerinde şemaları öğrenme sonucunu değerlendirmek için birçok farklı öğrenme algoritmaları için tek tip bir arayüz sağlar. Weka kolayca veri kümesi için geçerli olabilir öğrenme algoritmalarının uygulamalarından sağlar. Aynı zamanda, bu tür algoritmalar olarak transforme veri setleri için çeşitli araçlar içerir,

Weka tezgahı algoritmaları ve araçları önişleme verileri öğrenme state-of-the-art makinesinin topluluğudur. Hızla esnek şekillerde yeni veri setleri üzerinde var olan yöntemleri deneyebilirsiniz böylece tasarlanmıştır. Öğrenme algoritmaları çeşitli olarak kuyudan olarak, ön işleme araçları geniş bir yelpazede içerir. Kullanıcılarının farklı yöntemler karşılaştırın ve eldeki sorun için en uygun olanları tespit böylece Bu çeşitli ve kapsamlı araç seti ortak bir arayüz üzerinden erişilir. Şekil 1'de gösterildiği gibi Weka kullanımı en kolay yolu menü seçimi ve form doldurmayı kullanarak tüm tesislerinin erişim sağlar Explorer adlı grafik kullanıcı ara yüzü aracılığıyla hayata geçirilmektedir.



Şekil 3.1 Weka GUI Explorer

Weka bu işlevselliği kolay erişim için grafik kullanıcı arayüzleri ile birlikte, görselleştirme araçları ve veri analizi ve öngörü modelleme için algoritmaların bir koleksiyonu içerir. Weka avantajları şunlardır :

Tamamen böylece Java Programlama dili uygulanması ve GNU

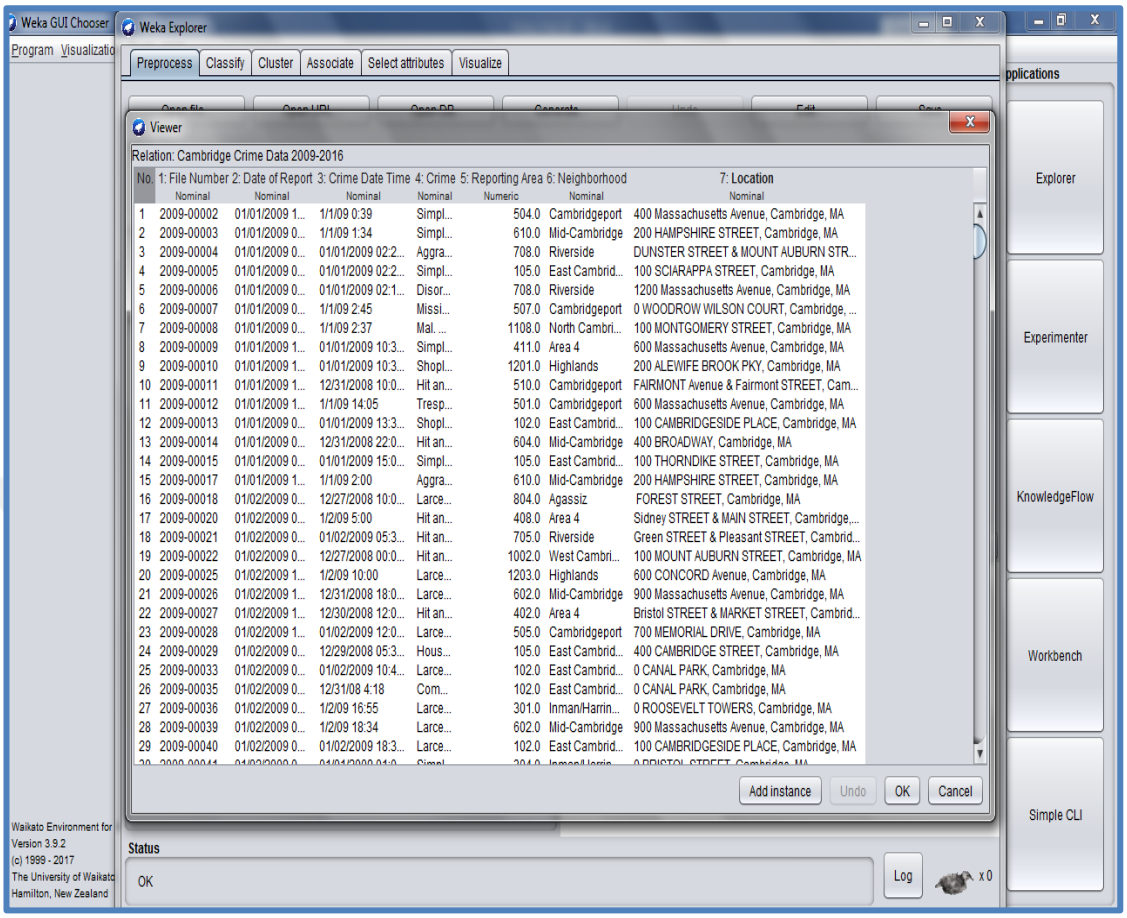
Genel Kamu Lisansı Taşınabilirlik altında Free kullanılabilirlik, hemen her modern bilgisayar platformu üzerinde çalışır.

Önişleme ve teknikleri modelleme verilerin kapsamlı bir koleksiyonu.

Grafik kullanıcı arayüzleri kullanım kolaylığı

Weka çeşitli standart veri madenciliği daha spesifik görevleri, veri ön işleme, kümeleme, sınıflandırma, regresyon, görselleştirme ve özellik seçimi destekler. Tüm veriler, her bir veri noktası normal olarak, sayısal veya nominal özellikler, fakat başka bir özelliğiyse türleri de desteklenmektedir özelliklerine (sabit bir sayısına göre tarif edilen tek bir düz dosya ya da bunlarla ilişkili olarak mevcut olduğu varsayımına dayanarak edilir). Weka Java Veritabanı Bağlantısı'nı kullanarak SQL veritabanlarına erişim sağlar ve bir veritabanı sorgusu tarafından döndürülen sonuçları işleyebilir. WEKA ana kullanıcı arayüzü Explorer, ama temelde aynı işlevi bileşen tabanlı Bilgi Akışı arayüz üzerinden ve komut satırından erişilebilir. Veri kümelerinin

bir koleksiyonu Wekas makine öğrenme algoritmaları olup öngörü performansının sistematik karşılaştırılmasını sağlar.

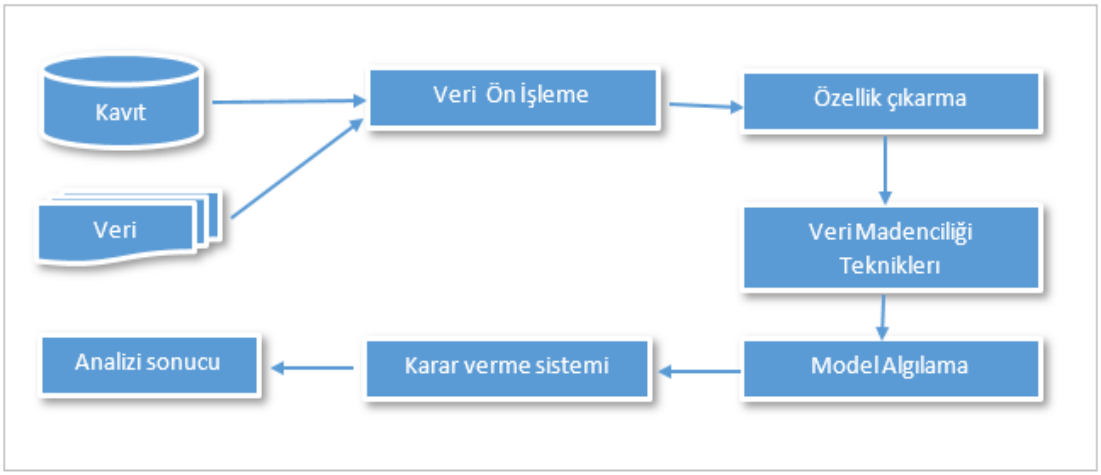


Şekil 3.2 WEKA'da kullanılan veriler

3.2 Metodoloji

Gelecekte belirli bir zamanda potansiyel olarak tehlikeli noktaları tahmin suç unsurları arasındaki bulgu ilişkiyi olduğuna son derece inanıyoruz. Bu nedenle, bizim önerilen yaklaşım suç, olay zamanında ve suç yerinin türüdür suçları verilerin üç ana unsurları ortaya koymaktır. Veri anlayışı, veri hazırlama, modelleme, değerlendirmeden oluşur. Kolluk veriler heterojen kaynaklardan toplanmaktadır. İşleme görevi, suç eğilimleri tespit o suç oranı azaltılabilir, böylece ileride suç faaliyetlerinde sıcak noktaları tahmin etmek amacıyla suç ve suç bölgesinin yazmak için yapılacak olan hangi suç kaydındaki veriler anlaşılır. Toplanan kolluk verileri birçok formatta olduğu için, veri ön işleme verimli istenen doğru sonucu elde etmek için veri kümesi kalitesini yükseltmek için yapılır. Model gerçekleştirilir ve sonuçların görsel görüntüleme araçları

kullanılarak grafik yapılır. Özellik seçimi, kümeleme, analiz, önceden haber verme ve daha sonra değerlendirme işlemi gerçekleştirmek için tasarlanmış olması gerekir.



Şekil 3.3 Analiz metodolojisi şeması

Sabıka kaydı: Biz telefon kayıtları gibi farklı bir kaynaktan kolluk tarafından toplanan veri setleri özellikle vurgu ile suç önleme çabalarında bu işin içinde veri madenciliği tekniklerinin uygulanabilirliğini araştırmak amacı, konum Facebook, twitter, blog, gözetleme plak gibi sosyal ağlar dayalı polis kayıtları, olay yeri inceleme sürecine mali işlem verileri. Bu deney için kullanılan veri kümesi gerçek ve otantik. Veri kümesi data.world web sitesinden elde edilir. Veri kümesi başlığı atlcrime suç (Atlanta-2009-2017 (APD)) içindedir. Biz uygulayacağımız bu model suç desenleri ayıklamak için bize yardımcı olabilir. Daha sonra, bu yapılar bu tür ilişki kuralları, madencilik sınıflandırılması ve özellikleri suç değerler temelinde sınıflandırmak, suç kayıtlarına küme gibi bazı veri madenciliği teknikleri uygulanacaktır. Bu veriler aşağıdaki ana özelliklere sahip hem suç ve suçlularla ilgili verileri içerir:

1. Dosya Numarası,
2. Raporun tarihi,
3. Suç Tarih Saat,
4. Suç,
5. Raporlama Alanı,
6. Mahalle,
7. Yer

File Number,Date of Report,Crime Date Time,Crime,Reporting Area,Neighborhood,Location
2009-00002,01/01/2009 12:39:00 AM,1/1/09 0:39,Simple Assault,504,Cambridgeport,"400 Massachusetts Avenue, Cambridge, MA"
2009-00003,01/01/2009 01:34:00 AM,1/1/09 1:34,Simple Assault,610,Mid-Cambridge,"200 HAMPSHIRE STREET, Cambridge, MA"
2009-00004,01/01/2009 01:43:00 AM,01/01/2009 02:20 - 02:35,Aggravated Assault,708,Riverside,"DUNSTER STREET & MOUNT AUBURN STREET, Cambridge, MA"
2009-00005,01/01/2009 02:27:00 AM,01/01/2009 02:20 - 02:45,Simple Assault,105,East Cambridge,"100 SCIARAPPA STREET, Cambridge, MA"
2009-00006,01/01/2009 02:34:00 AM,01/01/2009 02:15 - 02:35,Disorderly,708,Riverside,"1200 Massachusetts Avenue, Cambridge, MA"
2009-00007,01/01/2009 02:45:00 AM,1/1/09 2:45,Missing Person,507,Cambridgeport,"0 WOODROW WILSON COURT, Cambridge, MA"
2009-00008,01/01/2009 02:37:00 AM,1/1/09 2:37,Mal. Dest. Property,1108,North Cambridge,"100 MONTGOMERY STREET, Cambridge, MA"
2009-00009,01/01/2009 10:42:00 AM,01/01/2009 10:30 - 10:42,Simple Assault,411,Area 4,"600 Massachusetts Avenue, Cambridge, MA"
2009-00010,01/01/2009 10:47:00 AM,01/01/2009 10:35 - 10:47,Shoplifting,1201,Highlands,"200 ALEWIFE BROOK PKY, Cambridge, MA"
2009-00011,01/01/2009 12:37:00 PM,12/31/2008 10:00 - 01/01/2009 10:00,Hit and Run,510,Cambridgeport,"FAIRMONT Avenue & Fairmont STREET, Cambridge, MA"
2009-00012,01/01/2009 12:43:00 PM,1/1/09 14:05,Trespassing,501,Cambridgeport,"600 Massachusetts Avenue, Cambridge, MA"
2009-00013,01/01/2009 02:31:00 PM,01/01/2009 13:30 - 13:45,Shoplifting,102,East Cambridge,"100 CAMBRIDGESIDE PLACE, Cambridge, MA"
2009-00014,01/01/2009 02:35:00 PM,12/31/2008 22:00 - 01/01/2009 02:00,Hit and Run,604,Mid-Cambridge,"400 BROADWAY, Cambridge, MA"
2009-00015,01/01/2009 03:07:00 PM,01/01/2009 15:00 - 15:30,Simple Assault,105,East Cambridge,"100 THORNDIKE STREET, Cambridge, MA"
2009-00017,01/01/2009 11:52:00 PM,1/1/09 2:00,Aggravated Assault,610,Mid-Cambridge,"200 HAMPSHIRE STREET, Cambridge, MA"
2009-00018,01/02/2009 01:28:00 AM,12/27/2008 10:00 - 01/02/2009 01:10,Larceny from MV,804,Agassiz," FOREST STREET, Cambridge, MA"
2009-00020,01/02/2009 05:00:00 AM,1/2/09 5:00,Hit and Run,408,Area 4,"Sidney STREET & MAIN STREET, Cambridge, MA"
2009-00021,01/02/2009 05:33:00 AM,01/02/2009 05:30 - 05:32,Hit and Run,705,Riverside,"Green STREET & Pleasant STREET, Cambridge, MA"
2009-00022,01/02/2009 08:03:00 AM,12/27/2008 00:01 - 12/28/2008 13:46,Hit and Run,1002,West Cambridge,"100 MOUNT AUBURN STREET, Cambridge, MA"
2009-00025,01/02/2009 10:02:00 AM,1/2/09 10:00,Larceny (Misc),1203,Highlands,"600 CONCORD Avenue, Cambridge, MA"
2009-00026,01/02/2009 10:55:00 AM,12/31/2008 18:00 - 01/02/2009 11:00,Larceny from MV,602,Mid-Cambridge,"900 Massachusetts Avenue, Cambridge, MA"
2009-00027,01/02/2009 12:44:00 PM,12/30/2008 12:00 - 01/02/2009 12:30,Hit and Run,402,Area 4,"Bristol STREET & MARKET STREET, Cambridge, MA"
2009-00028,01/02/2009 12:11:00 PM,01/02/2009 12:04 - 12:05,Larceny from MV,505,Cambridgeport,"700 MEMORIAL DRIVE, Cambridge, MA"

Şekil 3.4 Suç veri kümesi

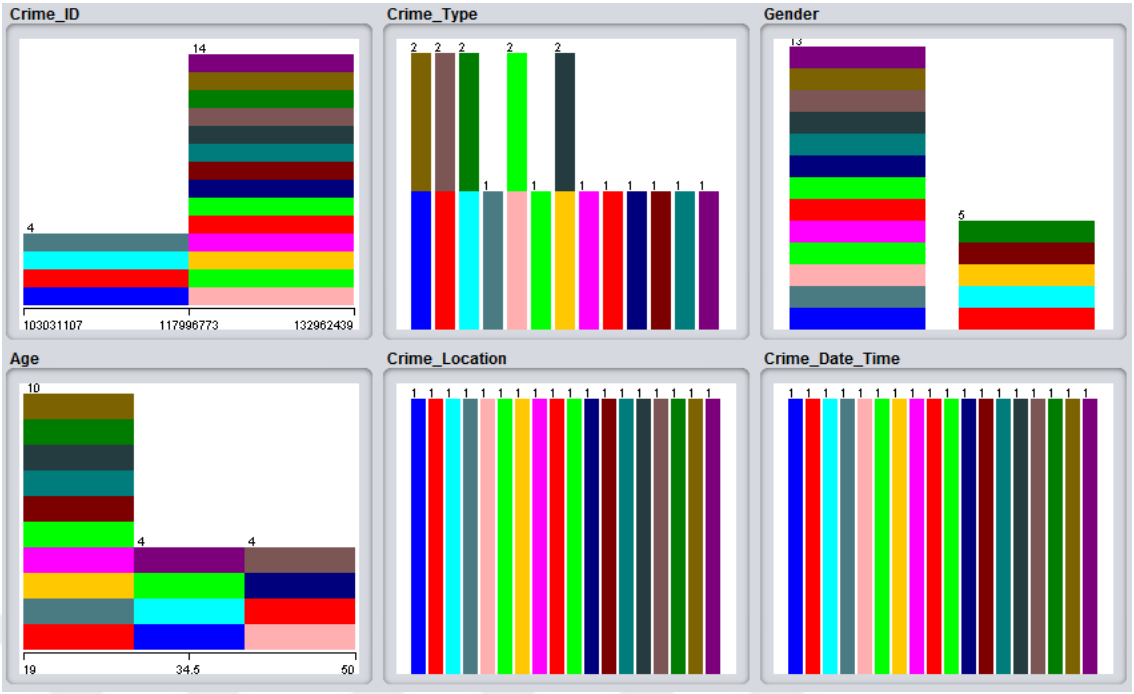
Veri Ön işleme: Veri ön işlem tekniği iyi kaliteli sonuçlar elde etmek için kullanılır. Ham veriler farklı bir biçimde veri ayıklama öncesinde önceden işlenir. Onlar veritabanları yanı sıra veri depolarından toplanmaktadır. Bu deney için kullanılan veri kümesi gerçek ve otantik. Deney için kullanılan veri kümesi bazı eksik değerleri içeren toplam 5001 durumlarda oluşur. Veri işleme gerçekleştirmek için, veri kalitesini artırmak için esastır. Veri ön işleme amacıyla kullanılır pratikte birkaç teknik vardır. Teknikler veri temizliği, entegrasyon, redüksiyon ve dönüştürme bulunmaktadır. Bir sınıflandırma algoritması uygulamadan önce genellikle bazı ön işleme veri kümesi üzerinde gerçekleştirilir. İlk adımda, veri azaltma sınıflandırma için hiçbir kritik bilgiler kaybetmeye çalışırken, bir veri kümesindeki en bilgilendirici özelliklerini seçerek gerçekleştirilir. Sadece altı nitelik 7 niteliklerin koleksiyonundan seçilir. Bu deneme için, manuel yöntem nitelik seçimi için seçildi. İkinci adımda, bazı haberler ayrıntılarının 'Cinsiyet, Yaş' adlı veri kümesi eklendi. Bunlar eklenen özellikler cezai cinsiyet değerine dayanmaktadır. Bu haber özelliklerini eklemek nedeni, tahmini gerçekleştirmek için amaç niteliği doğada itibari olmalıdır olduğunu.

Veri Temizleme: Atlanta Emniyet Müdürlüğü veri kümesinde cinsiyet ve yaş gibi bazı özelliklerde eksik değerler vardır. Ancak, eksik değerler içeren tüm özelliklerin anahtar özelliklerimizden olmadığını tespit ettik. Bu nedenle, onları temizlememiz gerekmedi. Çizelgedaki tüm anahtar özellikler, veri kümesindeki temizlenmiş değerlerle tamamlandı. Ayrıca, bu özelliklerde herhangi bir gürültülü veya tutarsız değer bulamadık.

Veri entegrasyonu: Veri setimiz için birkaç veri entegrasyonu adımı gerçekleştirdik. İlk olarak, farklı özellik adlandırmalarından kaçınmak için, suç veri kümesi için temel özellik adlarını şu şekilde birleştirdik : Suç Kimliği, Suç Türü, Cinsiyet, Yaş, Suç Yeri ve Suç Tarihi. Suç Yeri Atlanta veri kümesi mahalle özelliğini, Cambridge veri kümesi için Bölge özelliğini temsil eder. Madencilik çalışmamız, farklı ayrıntı düzeylerine ilişkin tarih ve saat bilgilerinin analiz edilmesini gerektirir. Bu nedenle, üç özellik daha oluşturmak için tarih ve saat suçu bilgilerini içeren Suç Tarihi özelliğini kullandık :

Veri azaltma: Suç veri kümesi için, veri azaltma uygulamamız gerekiyordu, özellik alt küme seçimini kullanarak boyutsal azaltma uyguladık. Örneğin, Atlanta suç veri kümesindeki mevcut 7 özellik arasında, sadece altı tanesini seçtik. Seçilen özellikler, madencilik amacımız için ilgili özellikler veya anahtar özelliklerdir (bkz. Çizelge 2.1). Veri kümesinden diğer tüm alakasız özellikleri kaldırdık. Öte yandan, örnek sayısı açısından veri azaltımı gerçekleştirdik. Atlanta suçları veri kümesinin bir dizi soygun örneği içerdiğini gözlemledik.

Veri Dönüşümü ve ayrıklaştırma: Veri entegrasyon sürecimizi farklı farklı değerlere sahip olarak bitirdik. Bu iki özellik değerinin çeşitliliğini azaltmanın gerekli olduğunu fark ettik. Böylece, değerlerini daha küçük gruba düşecek şekilde eşleştirerek her iki özelliğe de veri dönüşümünü uyguladık. Hedefimiz daha sık kalıp almak ve model doğruluğunu arttırmaktı. Suç türleri özelliği için, tür listesini altı yeni türe ayırarak en aza indiririz.



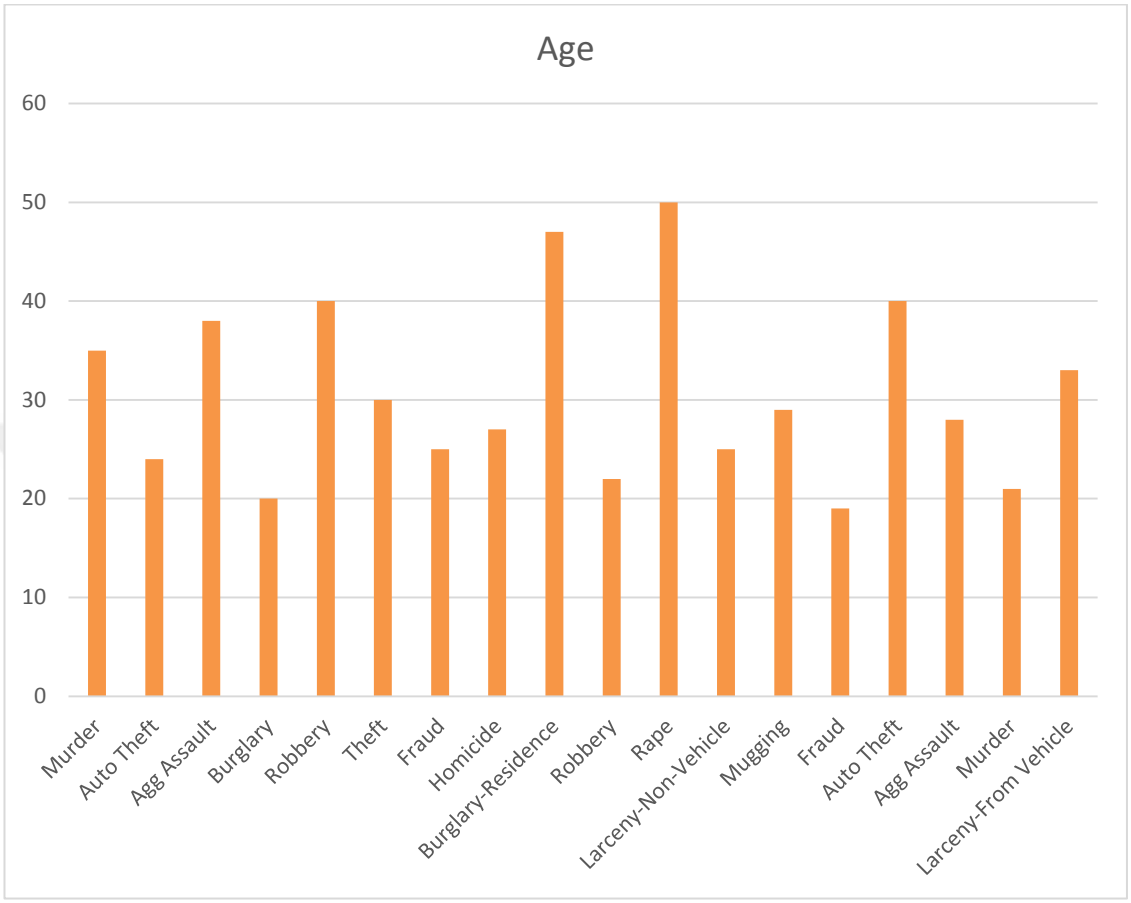
Şekil 3.5 Weka'daki grafik veri kümesinin gösterimi

Çizelge 3.1 Veri kümesi

Suç Kimliği	Suç Türü	Cinsiyet	Yaş	Suç Yeri	Suç Tarihi Saati
103040029	Cinayet	Erkek	35	610 Spring St Nw	31/10/2010
103040061	Oto Hırsızlığı	Kadın	24	850 Oak St Sw	31/10/2010
103040443	Agg Saldırısı	Kadın	38	1083 Euclid Ave Ne	31/10/2010
103031107	Hırsızlık	Erkek	20	370 Northside Dr Nw	30/10/2010
130331577	Soygun	Erkek	40	400 Formwalt St Sw	02/02/2013
130331607	Çalınması	Erkek	30	141 Auburn Ave Ne	02/02/2013
130331611	Dolandırıcılık	Kadın	25	1327 Sharon St Nw	02/02/2013
130331625	Cinayet	Erkek	27	3393 Peachtree Rd Ne	02/02/2013
130331697	Hırsızlık-Residence	Erkek	47	217 Hipp St Sw	02/02/2013
130331718	Soygun	Erkek	22	15 Wall St	02/02/2013
130331747	Soygun	Erkek	50	231 18th St Nw	02/02/2013

Özellik çıkarma : Varlık çıkarımı, suç paterni için temel bilgi sağlayacak şekilde belirli kalıpları tanımlama işlemidir. Varlık çıkarımı, kişinin kimliği, adresi, saati, yeri, tarihi, cinsiyeti, suç türü ve yaşı, belirli durumlar ile ilgili şüpheli tanımının otomatik olarak değerli temel bilgilerini çıkarmak için kullanılır. Potansiyel şüphelileri belirleme sürecidir. İlk veri madenciliği görevimiz, otomatik teknikler kullanılarak analiz edilmesi zor olan polis teşkilat raporlarından adlandırılmış tüzel kişiliklerin çıkarılmasını

içeriyordu. Atlanta Polis Departmanından (ADP) gürültülü olan ve birçok yazım hatası, yazım hatası ve dilbilgisi hatası içeren 20'den fazla rapor dosyasını rastgele seçtik.



Şekil 3.6 Suç analizi grafiği

4. VERİ MADENCİLİGİ GÖREVLERİ

Krishnamurthy ve Kumar, (2012) suç bilgilerinin veri madenciliği problemine nasıl dönüştürüleceğine bakacağız, böylece dedektiflerin suçları daha hızlı çözmesine yardımcı olabiliriz. Suç terminolojisinde bir kümenin coğrafi bir bölgedeki bir suç grubu ya da sıcak bir suç noktası olduğunu gördük. Oysa, veri madenciliği terminolojisinde bir küme benzer veri noktalarından oluşan bir grup olası bir suç modelidir. Dolayısıyla, uygun kümeler veya kümenin bir alt kümesi, suç modellerine birebir karşılık gelecektir. Bu nedenle, veri madenciliğindeki kümeleme algoritmaları, kendi aralarında benzer ancak verilerin geri kalanından farklı olan kayıt gruplarını tanımlama görevine eşdeğerdir. Bizim durumumuzda, bu kümelerin bazıları, bir veya aynı şüpheli grubu tarafından işlenen bir suç çılgınlığını tanımlamak için faydalı olacaktır. Bu bilgi göz önüne alındığında, bir sonraki zorluk en iyi kümelenmeyi sağlayan değişkenleri bulmaktır. Bu kümeler daha sonra alan uzmanlıklarını kullanarak detaya inmek için dedektiflere sunulacak. Suç kalıplarının otomatik tespiti, dedektiflerin önce suç çılgınlarına odaklanmalarına izin verir ve bu suçlardan birini çözmek tüm "çılgınlığı" çözme sonuçlarıyla sonuçlanır veya bazı durumlarda olay gruplarının bir çılgınlık olduğundan şüpheleniliyorsa, tam kanıt Suç olaylarının her birinden farklı bilgi parçalarından inşa edilebilir. Bugün çoğu, dedektiflerin genellikle bilgisayar veri analistlerinden ve kendi suç günlüklerinden aldıkları çoklu elektronik Çizelge raporlarının yardımıyla elle yapılır. Kümelenme tekniğini, sınıflandırma gibi herhangi bir denetimli teknik üzerinde kullanmayı tercih ediyoruz, çünkü suçlar doğada büyük ölçüde değişmektedir ve suç veri tabanı genellikle birkaç çözülmemiş suç içermektedir. Bu nedenle, mevcut ve bilinen çözülmüş suçlara dayanacak olan sınıflandırma tekniği, gelecekteki suçlar için iyi bir öngörücü kalite vermeyecektir. Ayrıca, internet tabanlı siber suçlar veya cep telefonu kullanan suçlar gibi zaman içinde meydana gelen suçların niteliği çok uzun zaman önce görülmedi.

Örneğin, bir suç sitesi şüphelinin siyah saçları olduğunu, bir sonraki olay / tanık şüphelinin orta yaşlı olduğunu ve üçüncüsü sol kolda dövme olduğunu ortaya koyuyor, hep birlikte bunlardan herhangi birinden çok daha eksiksiz bir resim verecek tek başına. Şüpheli bir suç paterni olmadan, dedektifin, farklı suç olaylarından bilgi bitlerinden tam resmi oluşturma olasılığı daha düşüktür. Böylece, gelecekte daha yeni ve bilinmeyen kalıpları tespit edebilmek için, kümeleme teknikleri daha iyi çalışır. Operasyonel zeka

ve veri madenciliği yoluyla suç analizi, bu nedenle erişilebilir verilerde basılan ve çok sayıda suç olayı tarafından üretilen belirli kalıpları tespit etmek istiyor. Bu paternler tekrarlanması muhtemel paternleri yansıtır ve bunların tanımlanması ve analizi, tekrarlarını önleyecek veya bozacak en uygun proaktif önlemleri gösterir. Birçok çalışma, suç analizi ve veri madenciliği arasındaki bu ittifakın değerini göstermektedir. Pazarlama alanında elde edilen başarılara dayanarak, veri madenciliği öncelikle finansal suç analizine, özellikle kredi kartı sahtekarlığına uygulandı. Bununla birlikte, ele geçirilen narkotiklerin kimyasal verilerindeki kalıpların keşfi ile organize suç ve siber suç veya uyuşturucu kaçakçılığı gibi diğer suç türlerine de uzanır.

Kaur, N. (2016) bununla birlikte, belirli uygulama alanlarındaki bu çalışmalar, ceza analizinde veri madenciliğinin kapsamını sınırlayacak ve adli ve kriminolojik bilgilerin enjeksiyonu arasında gerekli dengeyi tanımlamaya yardımcı olacak herhangi bir şema veya yansıma olduğunu göstermemektedir. A priori ve denetimsiz veri madenciliği tekniklerinin özgürlüğü. Bazı suç kalıplarını tespit etmek kolaydır. Örneğin, tespit eşikleri belirleyerek ve bir hedef bölgedeki suç oranlarını önceki haftalardaki oranlarla karşılaştırarak, eğilimler tespit edilebilir hale gelir. Diğer durumlarda, sismik kopyaları öngören modeller, benzer şekilde, suçların mekansal-zamansal dağılımını takip edecek şekilde uyarlanmıştır. Suç soruşturması, aranan kalıpların doğasını ve veri madenciliğinin sınırlarını anlamaya yardımcı olan özellikler sunar. Bir suç faaliyeti, zorunlu olarak, maddi izler şeklinde çevrilebilen, fakat aynı zamanda uzama, mekansal-zamansal özellikler veya aktiviteden kaynaklanan diğer herhangi bir unsurla çevrilebilen etkiler üretecektir.

Araştırmanın a mantığı, izlerden veri toplayarak, bunları analiz ederek, sonra olanları yeniden yapılandırmaya çalışarak ilerler. Bu süreç kusurludur, çünkü eksik verilere dayanır (tüm olaylar rapor edilmez), ilgili verilerin tamlığını garanti edemeyen bir koleksiyon (izler varsa, araştırmacılar tarafından algılanıp toplandıklarına dair bir garanti yoktur). Ek olarak, yeniden yapılanma, etkilerden nedenlere doğru ilerler kaçırma yoluyla, yaklaşık, belirsiz ve gözden geçirilebilir akıl yürütmede (birkaç hipotez gözlemlenen etkileri açıklayabilir).

4.1 Sınıflandırma

Estivill-Castro ve Lee, (2001) sınıflandırma, tahminlerin modellenmesi için bir teknik olarak veri madenciliğinin önemli özelliklerinden biridir. Doğada denetlenen bir sınıf tahmin tekniğidir. Bu teknik, yeterli sayıda eğitim örneğinin mevcut olması kaydıyla sınıflar için etiketi tahmin etme yeteneğine sahiptir. En yakın Komşular, Karar Ağacı, Naive Bayes ve Yapay Sinir Ağları gibi destek vektör makineleri de dahil olmak üzere çeşitli sınıflandırma algoritmaları mevcuttur. Tüm bu teknikler, bilinmeyen sınıf etiketini tahmin etmek için model kümesini keşfetmek amacıyla bir veri kümesine uygulanabilir. Sınıflandırmada veri seti, eğitim seti (bağımlı set) ve bir test seti (bağımsız set) olmak üzere iki gruba ayrılır. Sınıflandırma, önceki karar alma esasına göre bazı gizli ve gelecekteki kararlara örnekler vermek için kullanılır. Sınıflandırma, çeşitli suç varlıkları arasında karşılıklı özellikler bulur ve bunları, gönderenin yapısal özelliklerine ve dilsel örüntülerine göre e-posta spam'in kaynağını belirlemek için uygulanan önceden tanımlanmış sınıflar halinde düzenler. Genellikle suç eğilimlerini tahmin etmek için kullanılan sınıflandırma, suç varlıklarını tanımlamak için gereken süreyi azaltabilir. Ancak, teknik önceden tanımlanmış bir sınıflandırma şeması gerektirir. Sınıflandırma ayrıca makul derecede eksiksiz bir eğitim ve test verisi gerektirir, çünkü yüksek derecede eksik veri tahmin doğruluğunu sınırlar.

4.1.1 Karar Ağacı

Qayyum, S. (2018) karar ağacı öğrenimi, bir karar ağacını, bir öğeyle ilgili (dallarda temsil edilen) gözlemleri, öğenin hedef değeri (yapraklarda temsil edilen) ile ilgili sonuçlarla eşleştiren tahmini bir model olarak kullanır. İstatistik, veri madenciliği ve makine öğreniminde kullanılan öngörülü modelleme yaklaşımlarından biridir. Hedef değişkenin sonlu bir değer kümesi alabileceği ağaç modellerine sınıflandırma ağaçları denir ; bu ağaç yapılarında yapraklar sınıf etiketlerini ve dallar bu sınıf etiketlerine yol açan özelliklerin birleşimlerini temsil eder. Hedef değişkenin sürekli değerler alabileceği karar ağaçlarına regresyon ağaçları denir. Karar analizinde, bir karar ağacı, kararları ve karar almayı görsel ve açık bir şekilde temsil etmek için kullanılabilir. Veri madenciliğinde, bir karar ağacı verileri tanımlar (ancak sonuçta ortaya çıkan sınıflandırma ağacı, karar verme için bir girdi olabilir). Bir karar ağacı oluşturmak için Quinlan'ın önceki ID3 algoritmasının bir uzantısı olan C4.5 algoritması kullanılır. Ağacı oluşturmak için, düğümlerin belirlenmesinde entropi ölçümü kullanılır. Entropisi

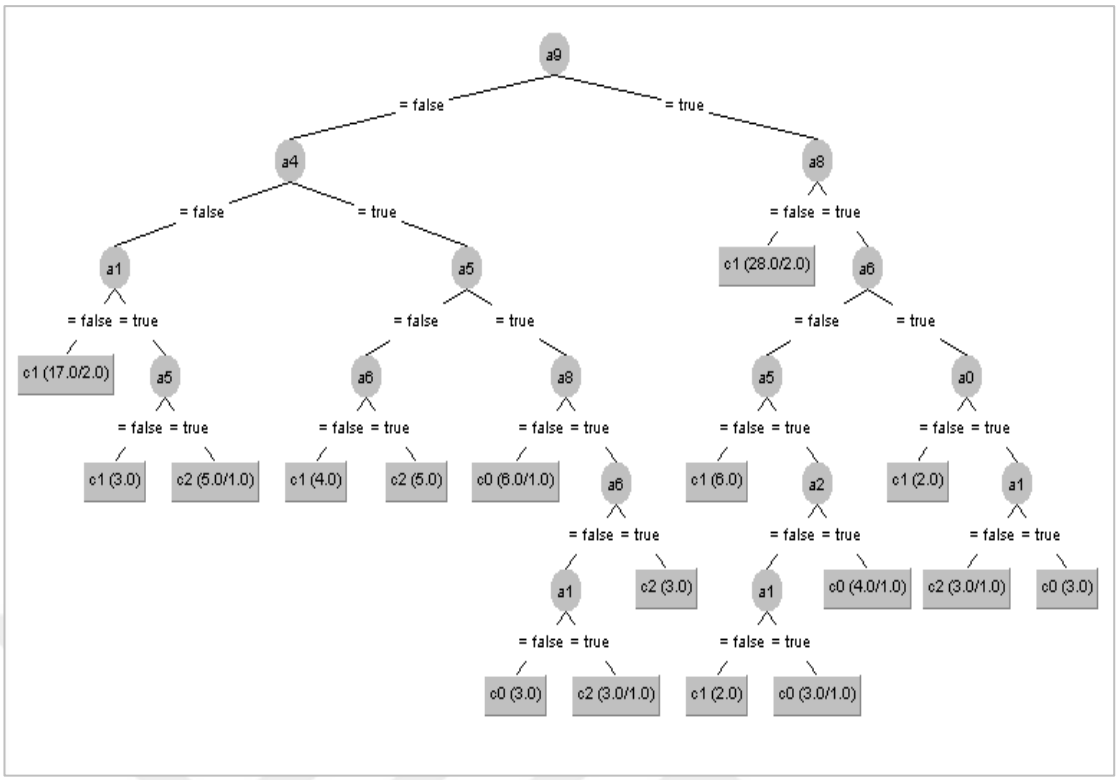
daha yüksek olan özellikler sonuçta daha belirsizliğe neden olduğundan, entropi sırasına göre seçildiler.

Ivan, Ahishakiye, Omulo, E. O., ve Taremwa (2017) bir ağaç, öznitelik değeri testine dayalı olarak kaynak kümesini alt kümelere bölerek "öğrenilebilir". Bu işlem, türetilmiş her alt kümede özyinelemeli bölümleme adı verilen özyinelemeli bir şekilde tekrarlanır. Özyinelemeli bölümleme veya özyinelemeli ikili bölme kullanılarak bölümlenmiş ve bölümlenmemiş alanlar için şekilde gösterilen örneklere bakın. Bir düğümdeki altkümenin hedef değişken ile aynı değere sahip olması veya bölünmesi durumunda özyineleme tamamlanır. Artık tahminlere değer katmıyor. Karar ağaçlarının yukarıdan aşağıya indüklenmesi süreci (TDIDT), açgözlü bir algoritma örneğidir ve veri için karar ağaçlarını öğrenmek için açık ara en yaygın stratejidir. Veri madenciliğinde, karar ağaçları, belirli bir veri kümesinin tanımlanmasına, sınıflandırılmasına ve genelleştirilmesine yardımcı olmak için matematiksel ve hesaplama tekniklerinin kombinasyonu olarak da tanımlanabilir. Veriler formun kayıtlarında gelir :

(x, Y) = (x1, x2, x3 ……, xk, Y). Bağımlı değişken olan Y, anlamaya, sınıflandırmaya veya genellemeye çalıştığımız hedef değişkendir. X vektörü, bu görev için kullanılan giriş değişkeni x1, x2, x3 vb.

Çizelge 4.1 Tabakalandırılmış çapraz geçerlilik

Doğru Sınıflandırılmış Örnekler	69	69 %
Yanlış Sınıflandırılmış Örnekler	31	31 %
Kappa istatistiği	0.4173	
Ortalama mutlak hata	0.226	
Karekök ortalama hata	0.4084	
Göreceli mutlak hata	61.9717 %	
Kök göreli kare hatası	95.9294 %	
Toplam Örnek Sayısı	100	



Şekil 4.1 Karar Ağacı

4.1.2 Naive Bayes

Naive Bayes algoritması, denetimli bir öğrenme yönteminin yanı sıra sınıflandırma için istatistiksel bir yöntemdir. Naive Bayes sınıflandırıcı, bir girdi verildiğinde tek bir çıktı sağlamak yerine tüm sınıfların kümesinin olasılık dağılımını veren olasılıklı bir sınıflandırıcıdır. Algoritma, bir haber makalesini en iyi uyduğu suç türüne göre sınıflandırır. Sınıflandırmadan ne elde ediyoruz ”demektedir. Bir suç belgesi D'nin belirli bir C sınıfına ait olma olasılığı nedir ? Naive Bayes sınıflandırıcı basit ve lojistik regresyon daha hızlı yakınsama vardır. Naive Bayes algoritması ile vandalizm, cinayet, soygun, hırsızlık, cinsel istismar, gang tecavüz, kundakçılık, silahlı soygun, otoyol, kapma vb. Onları bilinmeyen girdiler için test edebiliriz İle ilgili model bye eğitim suç verileri oluşturuyoruz. Modelin doğruluğunu test etmek için test verilerini uygularız. Ayrıca, sıfır frekans problemini giderir. Test sonuçları, Naive Bayes'in% 90'dan fazla doğruluk gösterdiğini, çünkü her kelimeyi jeton olarak sınıflandırdığını ve doğruluğu artıran “the”, “ve”, “of” vb. Bir kelime daha az veya 3 kereden az oluşmuşsa otomatik olarak sonlandırılır. Şekil 3.1, Naive Bayes algoritmasının örnek sahte kodunu göstermektedir.

Çizelge 4.2 Tabakalı çapraz doğrulama

Doğru Sınıflandırılmış Örnekler	61	61 %
Yanlış Sınıflandırılmış Örnekler	39	39 %
Kappa istatistiği	0.1697	
Ortalama mutlak hata	0.3129	
Karekök ortalama hata	0.4075	
Göreceli mutlak hata	85.7867 %	
Kök göreli kare hatası	95.7145 %	
Toplam Örnek Sayısı	100	

4.2 Kümeleme Tekniği

David ve Suruliandi (2017) bölüm kümeleme yöntemleri öncelikle K-ortalamaları, AK-modu ve Beklenti-Maksimizasyon algoritmaları olarak sınıflandırılır. Bölümleme yöntemi, belirli bir "n" nesnelerinin veri kümesinden "k" bölümlerini oluşturur. Ön işlemden sonra, operasyonel veriler nesneleri farklı kümeler olarak gruplamak için kümeleme tekniklerinden geçmektedir.

K-kümeleme algoritması anlamına gelir : K-araçları esas olarak kümeleri araçlarına göre bölümlemek için kullanılan algoritma anlamına gelir. Başlangıçta nesne sayısı gruplandırılır ve k kümesi olarak belirtilir. Ortalama değer, nesneler arasındaki ortalama mesafe olarak hesaplanır. Nesneleri bir gruptan diğerine taşıyarak bölümleri iyileştirmek için kullanılan yer değiştirme yinelemeli teknik. Yineleme sayısı yakınsama oluşana kadar yapılır.

Ak-modu algoritması: Ak-modu kümeleme algoritması, özellik tartım aşaması ve kümeleme aşaması gibi iki aşamalı bir işlemdir. Özellik tartım aşamasında, özelliklerin ağırlıkları her özellik için bilgi Kazanç Oranı (IGR) değeri kullanılarak hesaplanır. En büyük ağırlık değeri belirleyici bir özellik olarak alınır. İki kategorik öznitelik arasındaki mesafe, iki veri kaydı arasındaki fark benzerlik ölçümleri verdiği için hesaplanır. Analist, benzerlik ölçüsünün hesaplama sonucunun yardımıyla a eşik değerini belirlemiştir. Bu algoritma temel olarak kategorik özellikler için kullanılır.

Beklenti-Maksimizasyon algoritması: Beklenti-Maksimizasyon algoritması, her küme için parametre tahminlerini bulmak için kullanılabilen K-ortalama algoritmasının bir uzantısıdır. Tüm veriler parametrik olasılıksal dağılımın bir karışımıdır. Niteliklerin ağırlığı olasılık dağılımında ölçülür ve her bir nesne, nesneleri K-araçlarında ayrılmış

kümelere atamak yerine ağırlıklara göre kümelenmelidir. Parametre tahminlerini bulmak için, yinelemeli arıtma algoritmasının iki adımı kullan.

Kümeleme teknikleri, sınıflar arası benzerliği en aza indirmek veya en üst düzeye çıkarmak, suçları benzer şekilde taşıyan veya farklı çetelere ait gruplar arasında ayrım yapan şüphelileri belirlemek için veri nesnelerini benzer özelliklere göre sınıflar halinde gruplandırır. Bu tekniklerin, öğeleri atamak için önceden tanımlanmış bir dizi sınıfı yoktur. Bazı araştırmacılar, suç kayıtlarında kişi, kuruluş ve araç gibi farklı nesneleri otomatik olarak ilişkilendirmek için istatistik tabanlı konsept alanı algoritmasını kullanır. Benzer işlemleri tanımlamak için bağlantı analizi tekniklerini kullanarak, mali suçlar İcra Ağı AI Sistemi kara para aklama ve diğer mali suçların tespitini ve analizini desteklemek için Banka Gizlilik Yasası verilerinden yararlanır. Kümelenme suç olayları, suç analizinin önemli bir bölümünü otomatikleştirebilir, ancak genellikle gereken yüksek hesaplama yoğunluğu ile sınırlıdır.

Küme adı verilen her grup, kendi aralarında benzer ve diğer grupların nesnelerinden farklı nesnelerden oluşur. Verileri daha az kümeyle temsil etmek zorunlu olarak bazı ince ayrıntıları kaybeder (veri sıkıştırmasını berbat hale getirir), ancak basitleştirme sağlar. Birçok veri nesnesini birkaç kümeyle temsil eder ve bu nedenle verileri kümelerine göre modeller. Veri modelleme, kümelenmeyi matematik, istatistik ve sayısal analizden kaynaklanan tarihsel bir perspektife yerleştirir. Bir makine öğrenme perspektifinden, kümeler gizli kalıplara karşılık gelir, kümeler için araştırma denetimsiz öğrenmedir ve ortaya çıkan sistem bir veri kavramını temsil eder. Bu nedenle, kümeleme gizli bir veri kavramının denetimsiz öğrenilmesidir. Veri madenciliği, kümelenme analizine ek ciddi hesaplama gereksinimleri getiren büyük veritabanları ile ilgilenir. Bu zorluklar, aşağıda incelenen, geniş çapta uygulanabilir güçlü veri madenciliği kümeleme yöntemlerinin ortaya çıkmasına neden olmuştur. Bu çalışma alanına katkılarımızdan bazılarına bakacağız. Burada basit bir kümeleme örneği göstereceğiz. Çok basitleştirilmiş bir suç kaydı yapalım. Bir suç veri analisti, farklı verilere göre sıralanan bu verilere dayanan bir rapor kullanacaktır, genellikle ilk sıralama dedektifin deneyimine göre en önemli özellik olacaktır.

Çizelge 1'e basit bir suç listesi örneği ile bakıyoruz. Suç türü soygundur ve en önemli özellik olacaktır. 1'den 3'e kadar olan sıralar, şüpheli açıklamanın eşleştiği ve mağdur profilinin de benzer olduğu basit bir suç modelini göstermektedir. Buradaki amaç, veri

madenciliğini çok daha karmaşık kalıpları tespit etmek için kullanabilmemizdir çünkü gerçek hayatta suç için birçok özellik veya faktör vardır ve genellikle suçla ilgili kısmi bilgi mevcuttur. Genel bir durumda, bir bilgisayar veri analistinin bu kalıpları basit sorgulama ile tanımlaması kolay olacaktır. Bu nedenle, veri madenciliğini kullanan kümeleme tekniği, muazzam miktarda veriyle ve suç olayları hakkında gürültülü veya eksik verilerle başa çıkmak için kullanışlıdır. Burada k-araçları kümeleme tekniği kullandık, burada veri madenciliğinde en çok kullanılan tekniklerden biridir. Kümelemenin sınıflandırmaya kıyasla ana avantajı, uyum sağlama yeteneğidir ve farklı grupları ayıran yararlı özellikleri ayırmaya yardımcı olur. Kümeleme, aykırı değerleri algılamak için kullanılabilir. Kümeleme yöntemleri, bölümleme yöntemi, hiyerarşik yöntem, yoğunluk tabanlı yöntem, ızgara tabanlı yöntem, model tabanlı yöntem, kısıt tabanlı yöntem olarak sınıflandırılabilir.

4.3 Dernek kuralları

İlişkilendirme kuralı, bir veritabanında sık sık oluşan öğe kümelerini belirler ve kullanıcıların etkileşim geçmişinden bağlantı kuralları geliştirmek için ağ izinsiz giriş algılamasında kullanılan kurallar olarak bazı kalıplar sunar. Müfettişler, gelecekteki olası ağ saldırılarını algılamaya yardımcı olması için bu tekniği ağ davetsiz misafir profillerine de uygulayabilir. İlişkilendirme kuralı madenciliğine benzer olarak, sıralı model madenciliği, farklı zamanlarda gerçekleşen bir dizi işlem üzerinde sıkça oluşan öğe dizilerini bulur. Ağ saldırı tespitinde, bu yaklaşım zaman damgalı veriler arasındaki saldırı paternlerini tanımlayabilir. Gizli kalıpların gösterilmesi suç analizine fayda sağlar, ancak anlamlı sonuçlar elde etmek için zengin ve yüksek düzeyde yapılandırılmış veriler gerekir.

Hassani, Huang, Silva, ve Ghodsi (2016) ilişkilendirme kuralı, nitelikler arasındaki ilişkiyi tanımlar. İlişkilendirme kuralları şeklinde keşfedilen veri bilgisidir. İlişkilendirme kuralı madenciliği, belirli bir yerde suçun ilginç bir şekilde ortaya çıkmasını keşfetmenin bir yoludur. Sık öğe kümelerini bulur ve X Y ve / veya Y X için {X, Y} ilişkilendirme kuralları gibi sık kullanılan öğe kümelerinden oluşturur. Bir ilişkilendirme kuralının kendi destek değeri, güven değeri ve artışı vardır. Güven değeri ne kadar yüksek olursa ilişkilendirme kuralı o kadar güçlü olur. Yüksek bir artış X ve Y'nin birlikte varlığının sadece rastgele bir olay olmadığını, aynı zamanda ilişkilerinden kaynaklandığını gösterir.

(X –Y) kuralının desteği şu şekilde tanımlanır:

$$\text{Destek (x-y)} = \frac{\text{X ile sayı kayıtları}}{\text{Öğe kümelerindeki sayı kayıtları}}$$

(X–Y) kuralının desteği şu şekilde tanımlanır:

$$\text{Destek (x-y)} = \frac{\text{X ve Y ile sayı kayıtları}}{\text{Öğe kümelerindeki sayı kayıtları}}$$

(X –Y) kuralının güvenirliği şu şekilde tanımlanır:

$$\text{Güven (x-y)} = \frac{\text{X ve Y ile sayı kayıtları}}{\text{X ile sayı kayıtları}}$$

(X–Y) kuralının yükselmesi şu şekilde tanımlanır:

$$\text{Kaldırma (x-y)} = \frac{\text{Güven (x-y)}}{\text{Destek (x-y)}}$$

Şekil 4.2 Kümeleme analizi

Eğitim setinde model ve değerlendirme

Kümelenmiş Örnekler

0 37 (% 37)

1 63 (% 63)

Apriori algoritması

Minimum destek : 0.15 (15 örnek)

Minimum metrik <güven> 0,9

Gerçekleştirilen döngü sayısı : 17

Oluşturulan büyük öğe kümeleri :

Büyük parçaların set büyüklüğü L (1) : 23

Büyük parça setinin büyüklüğü L (2) : 201

Büyük parçaların set büyüklüğü L (3): 341

Büyük parçaların set büyüklüğü L (4) : 28

Bulunan en iyi kurallar :

1. a5 = yanlış a6 = yanlış 23 ==> sınıf = c1 23 <conf: (1)> asansör: (1.61) lev: (0.09) [8] dönş.: (8.74)

2. a3 = doğru a8 = yanlış a9 = doğru 18 ==> sınıf = c1 18 <conf : (1)> asansör: (1.61) lev: (0.07) [6] dönş: (6.84)

3. a5 = yanlış a6 = yanlış a9 = doğru 15 ==> sınıf = c1 15 <conf : (1)> asansör: (1.61) lev: (0.06) [5] dönş.: (5.7)

4. a2 = yanlış a8 = yanlış a9 = doğru 19 ==> sınıf = c1 18 <conf: (0.95)> asansör: (1.53) lev: (0.06) [6] dönş.: (3.61)

5. a2 = yanlış a6 = yanlış a8 = yanlış 18 ==> sınıf = c1 17 <conf: (0.94)> asansör: (1.52) lev: (0.06) [5] dönş.: (3.42)

6. a0 = doğru a6 = yanlış a7 = doğru 17 ==> a4 = yanlış 16 <conf: (0.94)> asansör: (1.74) lev: (0.07) [6] dönş.: (3.91)

7. a6 = yanlış a8 = yanlış a9 = doğru 17 ==> sınıf = c1 16 <conf: (0.94)> asansör: (1.52) lev: (0.05) [5] dönş.: (3.23)

8. a0 = doğru a8 = yanlış a9 = doğru 16 ==> sınıf = c1 15 <konf: (0.94)> asansör: (1.51) lev: (0.05) [5] dönş.: (3.04)

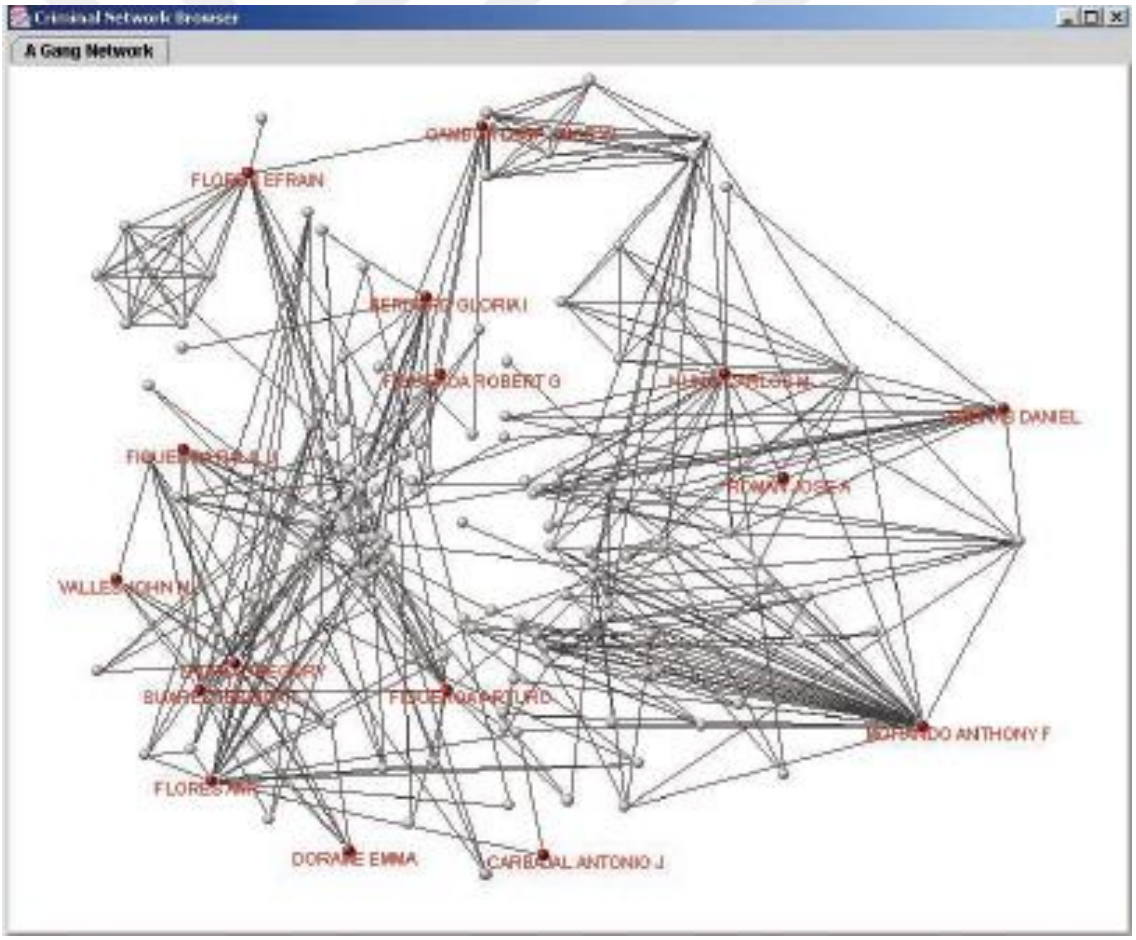
9. a8 = yanlış a9 = doğru 28 ==> sınıf = c1 26 <konf: (0.93)> asansör: (1.5) lev: (0.09) [8] dönş.: (3.55)

10. a2 = yanlış a6 = yanlış a9 = gerçek 20 ==> sınıf = c1 18 <conf: (0.9)> asansör: (1.45) lev: (0.06) [5] dönş.: (2.53)

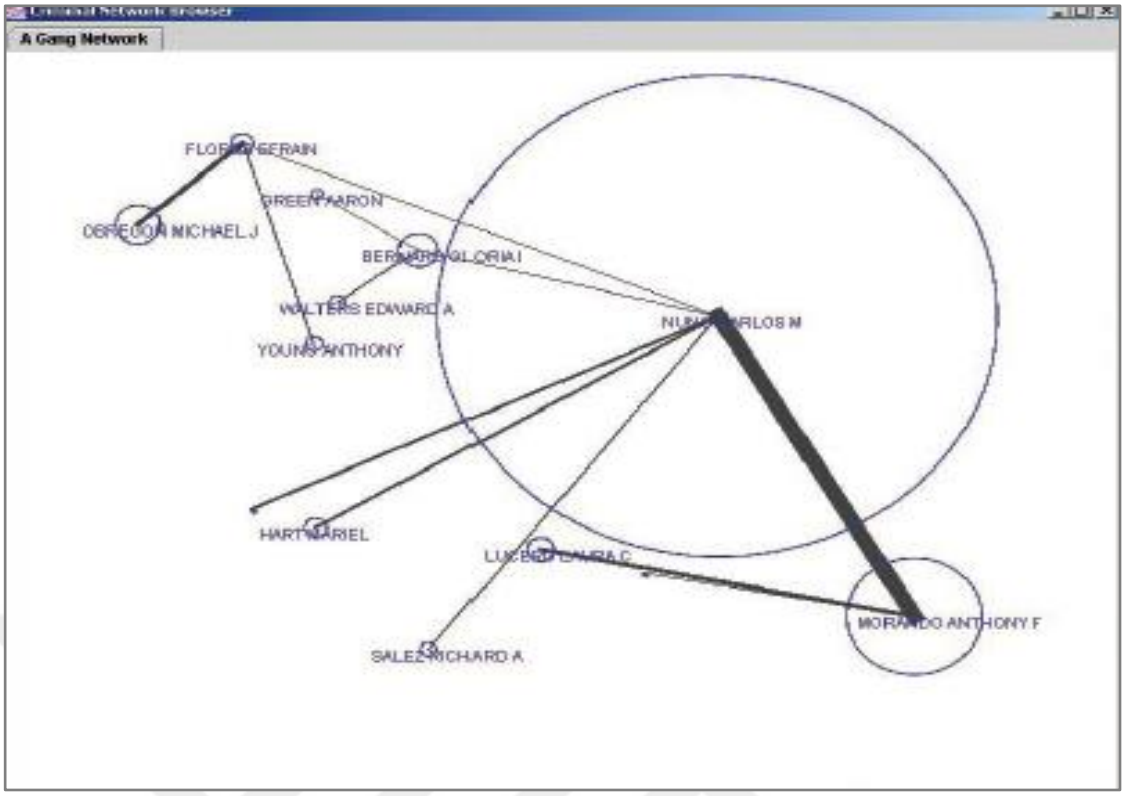
Örüntü Tespiti : Bu, suç eğilimlerini ve eğilimlerini tanımlamamız gereken suç modellerinin tanımlanması aşamasıdır. Suç modelini bulmak için, veritabanını taramak için çok verimli bir algoritma olan kümelemeyi kullanabiliriz. Mekansal paternler, şüpheli bilgiler ve operasyon yönteminin özelliklerini içeren suç verileri, araştırmacılara potansiyel şüphelileri tanımlamak için bir araç sağlamak için tutuklama alanları ile birleştirilebilir. Müfettişler sıcak noktaları ve söz konusu suçtan daha önce tutuklanan kişileri bu yerler etrafında hızlı bir şekilde bulabilirler.

Şüpheliler genellikle polis memurlarına yanlış adlar, adresler veya yaş verir ve bu nedenle birden fazla veritabanı girişine sahiptir ve bu da memurların şüphelinin gerçek kimliğini belirlemesini ve o kişiyle ilgili geçmiş olayları ilişkilendirmesini zorlaştırır. İkinci veri madenciliği görevimiz, Atlanta Emniyet Müdürlüğü'nün (APD) veritabanından kimlik, tip suç, cinsiyet, yaş, yer ve tarih gibi bilgileri içeren aldatıcı suç

kimliklerinin otomatik olarak tespit edilmesini içeriyordu. Bir vaka çalışmasında geliştirdiğimiz suç kimliği aldatma taksonomisine dayanarak, bir suçlunun kimliğini temsil etmek için suç türü, kimlik, adres ve sosyal güvenlik numarasını seçtik ve daha az güvenilir olan diğer alanları göz ardı ettik. Yöntemimiz, her kayıt çiftinin ilgili alanlarındaki değerleri karşılaştırmak için dize karşılaştırıcısı kullandı. Karşılaştırıcılar, 0 ile 1 arasındaki benzerlik değerlerini ölçer ve seçilen dört alan üzerinde Öklid vektör normu olarak iki kayıt arasındaki genel benzerlik ölçüsünü hesaplar. Bir Öklid vektör normu, kareli benzerlik ölçümlerinin toplamının kare köküdür ve ayrıca 0 ile 1 arasında normalleştirilmiştir. Eğitim aşamasında, aldatıcı ve aldatıcı olmayan kayıtları birbirinden ayıran 0.00 ile 1.00 arasında değişen eşik değerleri denedik. Eğitim için verilerin üçte ikisini ve test için kalanları kullanarak bir ayırma doğrulama yöntemi kullandık



Şekil 4.3 Desen Algılama



Şekil 4.3 Desen Algılama (devam)

Suç tahmini: David ve Suruliandi, (2017) suçun öngörülmesi için, geliştirilen her sistem veri madenciliği teknikleri yardımıyla kolluk kuvvetleri farklı algoritmaları takip etmektedir. Polis yetkilileri on yıllardan beri suçluların ve suçluların tarihini göz önünde bulundurarak bilgi ve geçmiş deneyimlerini kullanıyor, "geleceğin en iyi öngörücüsü geçmiş" suç tahminlerini geliştirmenin yaygın bir yöntemidir. Gözlemlenen geçmiş değerlerin davranışını tanımlayan matematiksel modeller, gelecekteki suç eğilimlerinin zaman serisi analizini geleceğe yansıtarak gelecekteki suç eğilimlerini tahmin etmek için kullanılabilir.

Almanie, Mirza ve Lor (2015) modelleme, değişkenlerin nedensel dizisini ve etkileşimlerin tahminini tanımlamayı içerir. Herhangi bir öngörücü model Endeavour'un belirli bağımsız (öngörücü) değişken ile bağımlı bir değişken (yani öngörülecek kriter) arasında bir ilişki göstermesi. Tahminlere rehberlik etmek için kullanılan çeşitli veri kaynakları ve yöntemler arasında suç istatistikleri ; uzmanlar, uygulayıcılar ve genel halk anketleri ; Edebiyat yorumları ; senaryo yazımı : ve gelecekteki suç eğilimlerini tahmin eden istatistiksel (zaman serisi) modeller.

Çizelge 4.3 Suç veri analiz yazılımı

Yazılım	**Açıklama**
XLMiner	Excel için Veri Madenciliği eklentisi
Zoom'un Görünümü	Eklenti raporlama çözümü
TIMi paketi	Akıllı Madencilik makinesi, yüksek ölçeklenebilirlik, hız, YG ve tahmin doğruluğu ile tahmin, segmentasyon ve veri hazırlama için bağımsız, otomatik, kullanıcı dostu bir GUI araçları ailesi
Synapse	Sinir ağları ve diğer uyarlanabilir sistemler için, model oluşturma ve eğitim yoluyla veri alma ve önişlemeden değerlendirme ve konuşlandırmaya kadar tüm geliştirme döngüsünü destekleyen bir geliştirme ortamı; .NET bileşenleri olarak konuşlandırmaya izin verir.
İstatistik Veri madencisi	Kapsamlı, entegre istatistiksel veri analizi, grafikler, veri tabanı yönetimi ve uygulama geliştirme sistemi.
SAS Enterprise Miner	SEMMA (Örnekleme, Keşfetme, Değiştirme, Modelleme, Değerlendirme) sürecine kullanıcı dostu bir GUI ön ucu sağlayan entegre bir paket olan Enterprise Miner
Salford Sistemi Veri Madenciliği Paketi	CART Karar Ağaçları. MARS tahmini modelleme, otomatik regresyon, TreeNet sınıflandırma ve regresyon, veri erişimi, hazırlama, temizleme ve raporlama modülleri, RandomForest tahmini modelleme, kümeleme ve anomali tespiti.
Previa	Sınıflandırma ve tahmin için ürün ailesi
Öngörülen Veriler	Tahminli Dinamiks'ten
Madencilik Süiti	Grafiksel ve istatistiksel veri analizini sinir ağları, kümeleme, bulanık sistemler ve genetik algoritmalar dahil olmak üzere modelleme algoritmalarıyla bütünleştirir.
Rap Analyst(tm)	Dinamik tahmin modelleri oluşturmak, yeni ve geçmiş veriler arasındaki ilişkileri ortaya çıkarmak için gelişmiş yapay zeka kullanır.
Powerhouse Veri Madenciliği yazılımı	Dorian Pyle'ın veri analizinde bilgi teorisini kullanma fikirlerine dayanan tahmin ve kümeleme modellemesi için.
Pentaho	Weka tabanlı raporlama, analiz, gösterge Çizelgeları, veri entegrasyonu ve veri madenciliği de dahil olmak üzere açık kaynaklı BI paketi

Oracle Veri Madenciliği (ODM)	Müşterilerin eyleme geçirilebilir öngörülebilir bilgiler üretmelerini ve entegre iş zekası uygulamaları oluşturmalarını sağlar.
Microsoft SQL Server	Microsoft BI platformuna sorunsuz bir şekilde entegre edilen ve herhangi bir uygulamaya genişletilebilen, sezgisel veri madenciliği yoluyla bilinçli analizlerle bilinçli kararlar verir.
Bilgi Madencisi (Yx)	Excel için, tahmin ve açıklayıcı modeller oluşturmak üzere Microsoft Excel'de depolanan verilerle çalışan bir bilgi madenciliği aracı. (MacOS, Excel 2004 veya üstü).
GMDH Shell	Tahminli modelleme ve veri madenciliği için gelişmiş ama kullanımı kolay bir araç.
Adil Isaac Model Oluşturucu	Analitik modeller geliştirmek ve dağıtmak için yazılım platformu, veri analizi, karar ağacı ve tahmine dayalı model oluşturma, karar optimizasyonu, iş kuralları yönetimi ve açık platform dağıtımını içerir.
Angoss bilgi stüdyosu	Kapsamlı bir veri madenciliği ve tahmine dayalı modelleme araçları paketi; SAS ve diğer önemli istatistiksel araçlarla birlikte çalışabilirlik.
II Karıncalar modeli oluşturucu	Regresyon, sınıflandırma ve eğilim modellerini içeren bir masaüstü tahmini analitik modelleme aracı.
II Karıncalar Tahmincisi	Kurumsal veritabanlarına karşı öngörülü modelleri dağıtmak için yüksek hızlı bir puanlama motoru.

Çizelge 4.3 Suç veri analiz yazılımı (demm)

Suç, yasa ve düzenden bir adım önde olmaya devam ettiğinden, sık sık geleneksel yöntemler ve araçlar, cihazlar ve teknoloji yapmak yetersiz olduğundan, polis departmanı suçla mücadele çabalarını artırıyor ve analiz edecek, tespit edecek sofistike yeni nesil yazılımları kullanarak daha akıllı hale geliyor suçları tahmin edebilir ve internetten gelen veriler, telefon görüşmeleri, parasal işlemler, trafik hareketleri, devriyeler de dahil olmak üzere suç kayıtlarının hacmini saniyeler içinde derleyebilir.

5. ARAŞTIRMA BULGULARI

Önerilen sistem jeo-uzamsal çizim ile birlikte kullanılmaktadır. Suç analisti belirli bir coğrafyadan bir zaman aralığı ve bir veya daha fazla suç türü seçebilir ve sonucu grafiksel olarak görüntüleyebilir. Bu gruptan, kullanıcı tüm seti veya ilgilenilen bölgeyi seçebilir. Ortaya çıkan veri kümesi, veri madenciliği işleme için giriş kaynağı haline gelir. Bu kayıtlar, önceden belirlenmiş niteliklere ve ağırlıklara göre kümelenir. Elde edilen bu kümeler jeo-uzamsal grafik üzerinde çizilir. Her grup için efsane, gruba dahil edilen önemli sayıda suç olayının yanı sıra grubu karakterize eden önemli nitelikleri de sağlar. Bu bilgi, dedektifin öngörülen suç kümelerini incelerken bakması için faydalıdır Kümelenme sürecinde mahkeme kararıyla ilgili hiçbir bilgi kullanılmamıştır.

Tespit edilen suç kalıpları için elde ettiğimiz sonuçları, bu suç olaylarına ilişkin mahkeme kararlarına bakarak şüpheliler hakkındaki suçlamaların kabul edilip edilmediğini doğruladık. Sonuç olarak, suçun demografisi gibi önemli nitelikler veya özellik veya suç değişkeni olarak ölçülen suç olayı verileri (bu suçların bazılarının sistemde halihazırda mevcut mahkeme kararları/kararları mevcuttu), şüpheli, mağdur vb. Daha sonra, suç yayılmalarının olası suç kalıplarını içeren suç gruplarını (veri madenciliği terminolojisindeki kümeler) bulmak için tartım tekniğimize dayanarak suçları kümelendiriyoruz. Suç kalıplarının jeo-uzamsal planı ve bu grupları ölçmek için önemli nitelikler, artık suç çılgınlarını tanımlamak için ilgili olmayan yüzlerce suç olayı listesinden veya önceden belirlenmiş bir türden çok daha kolay bir görevi olan dedektiflere sunulmaktadır sipariş. Bu, bu çalışmada deneme amacıyla oluşturulan yazılım prototipinin tasarım sürecini ve uygulamasını açıklar. Kullanılan veri seti, seçilen programlama stratejileri ve test süreci açısından sistemin uygulanması.

Bu aşamada, Weka'dan elde edilen sonuçları göstermek için aşağıdaki operatörler kullanılır.

• Model doğruluğu ve sınıflandırma hatası performans operatörleri tarafından hesaplanır,

• Günlük operatörü performans raporunu kaydetmek ve kaydetmek için kullanılır,

• Optimizasyon aykırı Tespit operatörü parametrelerinin etkisini değerlendirmek için bir doğruluk ve sınıflandırma hatası karşılaştırması kullanılır,

• Yeni çerçeveye dayalı suçların analizi.

Şekil 3.6, yeni şov çerçeve şemasını ayrıntılarıyla göstermektedir. Çizelge 1'deki sonuçlar arasındaki karşılaştırma, kümelerin sayısı benzer olduğunda, Outlier Detection operatörünün parametrelerini optimize ettikten sonra, sınıflandırma doğruluğunun arttığını ve sınıflandırma hatasının azaldığını ve sonuç olarak elde edilen uygunluk fonksiyonunun optimize edildiğini göstermektedir. Deneyde, suç veri kümesi Çizelge 2.1 üzerinde Naive Bayesian ve Karar Ağacı algoritmaları arasında bir karşılaştırma gerçekleştirilmiştir. Deney sırasında, önceden işlenmiş veri kümesi WEKA girişi 10 için standart dosya türünde. ARFF ve CSV dosyasına dönüştürülmüştür. Hem Naive Bayesian hem de Karar Ağacı algoritmaları için ayrı ayrı deneyde giriş veri kümesine çapraz doğrulama uygulandı. Naive Bayesian ve Karar Ağacı için 10 kat çapraz doğrulamanın doğruluğu sırasıyla % 61 ve % 69'dur.

Çizelge 5.1 Naive Bayesian Kullanan Karışıklık

	Düşük	**Yüksek**	**Orta**	
Düşük	4	14	3	21
Yüksek	5	54	3	62
Orta	3	11	3	17

Çizelge 5.2 Naive Bayesian Sınıfına Göre Doğruluk

	TP Oranı	**FP Oranı**	**Hassas**	**Hatırlama**	**F-Tedbir**	**MCC**	**ROC Alanı**	**PRC Alanı**	**Sınıf**
	0,190	0,101	0,333	0,190	0,242	0,112	0,659	0,300	c0
	0,871	0,658	0,684	0,871	0,766	0,254	0,756	0,850	c1
	0,176	0,072	0,333	0,176	0,231	0,137	0,720	0,360	c2
W. Avg	0,610	0,441	0,550	0,610	0,565	0,204	0,729	0,651	

Çizelge 5.3 Karar Ağacı Kullanarak Karışıklık Matrisi

	Düşük	**Yüksek**	**Orta**	
Düşük	10	8	3	21
Yüksek	9	52	1	62
Orta	6	4	7	17

Çizelge 5.4 Sınıf Karar Ağacına Göre Doğruluk

	TP Oranı	FP Oranı	Hassas	Hatırlama	F-Tedbir	MCC	ROC Alanı	PRC Alanı	Sınıf
	0,476	0,190	0,400	0,476	0,435	0,269	0,667	0,359	c0
	0,839	0,316	0,813	0,839	0,825	0,529	0,815	0,838	c1
	0,412	0,048	0,636	0,412	0,500	0,436	0,802	0,499	c2
W.Avg	0,690	0,244	0,696	0,690	0,688	0,459	0,782	0,680	

Daha yüksek performans. Ayrıca, Naive Bayesian ve Karar Ağacı için karışıklık matrisleri **Çizelge 3.5 ve Çizelge 3.7 '**de gösterilmektedir. **Şekil 4.2**, iki algoritma arasındaki karşılaştırmayı göstermektedir.

Çizelge 3.3 Naive Bayesian algoritması kullanılarak düşük, yüksek ve orta sınıfların sınıflandırılmasını göstermektedir. Karışıklık matrisinden elde edilen sonuç, aşağıdaki her sınıf için tartışılmaktadır :

Düşük sınıfında sınıflandırılan 21 ürün.

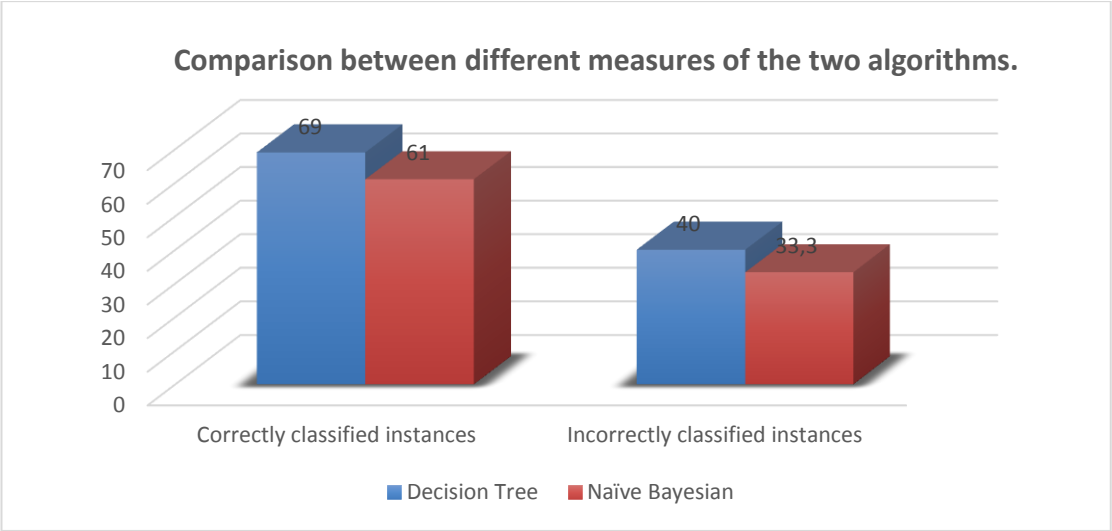
- Bunlardan 4'ü Düşük Sınıf olarak doğru bir şekilde sınıflandırılmıştır
- Bunlardan 14'ü yanlış Yüksek sınıf olarak sınıflandırılmıştır
- Bunlardan 3 tanesi yanlış bir şekilde Medium sınıfında sınıflandırılmıştır.

Düşük sınıfında sınıflandırılmış 62 öğedir.

- Bunlardan 5'i Düşük sınıfında doğru bir şekilde sınıflandırılmıştır
- Bunlardan 54'ü yanlış Yüksek sınıf olarak sınıflandırılmıştır
- Bunlardan 3 tanesi yanlış bir şekilde Medium sınıfında sınıflandırılmıştır.

Düşük sınıfında sınıflandırılmış 17 öğedir.

- Bunlardan 3'ü Düşük Sınıf olarak doğru sınıflandırılmıştır
- Bunlardan 11'i yanlış Yüksek sınıf olarak sınıflandırılmıştır
- Bunlardan 3 tanesi yanlış bir şekilde Medium sınıfında sınıflandırılmıştır.



Şekil 5.1 İki algoritmanın farklı ölçümleri arasındaki karşılaştırma.

Çizelge 5.5 Her iki Algoritma için doğruluk, yanlış sınıflandırılmış örnekler, Kesinlik, Geri Çağırma ve F-ölçümü

Yöntem	Doğruluk (Doğru sınıflandırılmış örnekler)	Yanlış sınıflandırılmış örnekler	Hassas	Hatırlama	F-Tedbir
Karar ağacı	69%	31 %	0,400	0,476	0,435
Naive Bayesian	61%	39%	0,333	0,190	0,242

Benzer şekilde, Çizelge 3.3 Karar Ağacı algoritması kullanılarak düşük, yüksek ve orta sınıfların sınıflandırılmasını gösterir. Karışıklık matrisinin sonucu aşağıdaki her sınıf için tartışılmaktadır :

Düşük sınıfında sınıflandırılan 21 ürün.

• Sınıfta doğru olarak sınıflandırılmış 10 madde vardır Düşük

• Yanlış Sınıf Yüksek olarak sınıflandırılmış 8 madde var

• Orta sınıfta yanlış sınıflandırılmış 3 madde var.

Düşük sınıfında sınıflandırılmış 62 öğedir.

• Sınıfta doğru sınıflandırılmış 9 madde vardır.

• Yanlış High Sınıfında sınıflandırılmış 52 madde var

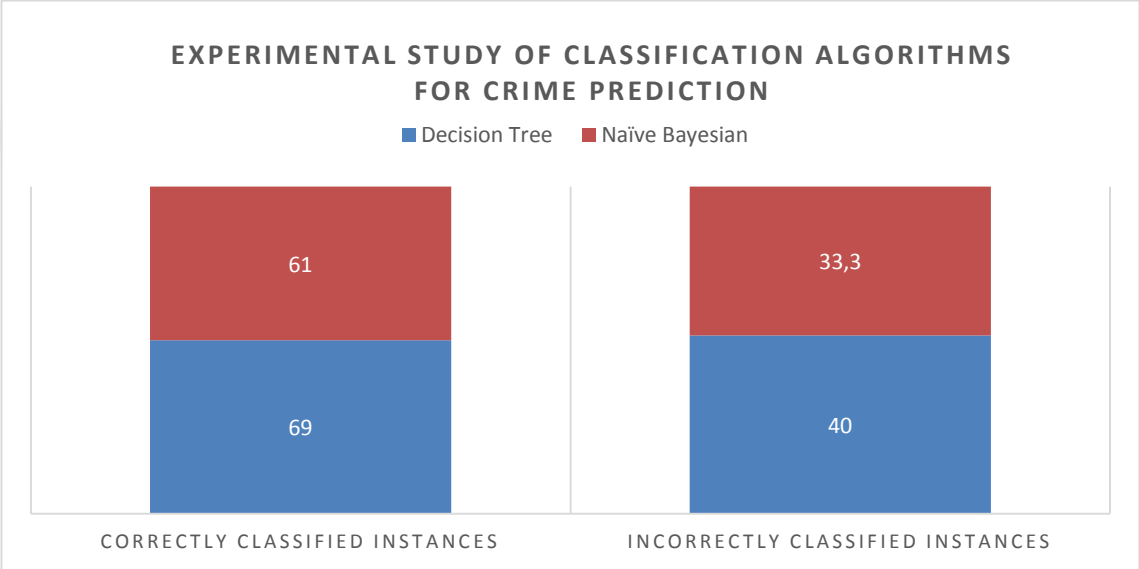
• Orta sınıfta yanlış sınıflandırılmış 1 madde var

Low sınıfında sınıflandırılmış 17 öğedir.

• Düşük sınıfa doğru sınıflandırılmış 6 madde var

• Yanlış Yüksek olarak sınıflandırılmış 4 madde var

• Orta sınıfta yanlış sınıflandırılmış 7 madde vardı



Şekil 5.2 Doğru ve yanlış sınıflandırılmış örneklerin yüzde karşılaştırması.

Karışıklık matrislerinin sonuçları ve yukarıdaki açıklaması, Karar Ağacı'nın Naive Bayesian'dan daha iyi performans gösterdiğini açıkça göstermektedir. Karar Ağacı, Düşük, Orta ve Yüksek olmak üzere tüm sınıfları tahmin etmede daha iyi performans gösterdi. Çizelge 4, deneyde kullanılan her iki algoritma için Doğruluk (doğru sınıflandırılmış örnekler), yanlış sınıflandırılmış örnekler, Hassasiyet, Geri Çağırma ve F-ölçümünü göstermektedir. Karar Ağacı için Doğruluk, kesinlik ve Geri Çağırma değerleri neredeyse aynıdır. Oysa Naive Bayesian'ta Doğruluk ve Hatırlama neredeyse aynıdır, ancak hassasiyetin değeri biraz daha azdır. İki algoritma için doğru sınıflandırılmış örneklerin yüzde karşılaştırması Şekil 10'da gösterilmektedir.

6. SONUÇ VE GELECEK ÇALIŞMALAR

Suç veri setini analiz etmenin mükemmel sonuçlarıyla ortaya çıkan veri madenciliği için kabul edilebilir bir modeldir; modeli oluşturmak ve test etmek için kullanılabilecek çok büyük tarihsel veriler gerektirir. Bu çalışmada kullanılan 5000'den fazla suç kaydı tahmin verebilir ve kabul edilebilir bir modele yol açabilir. Toplanan suç veri kümesini önceden işlemek ve analiz etmek için WEKA ve Excel yazılımı kullanılmıştır. Her şeyden önce toplanan veriler, WEKA'daki Denetlenmeyen öznitelikte bulunan Sayısal-Nominal'i kullanarak veri kümesini sayısaldan nominal değere dönüştürmek için ön işlemden geçirildi ve veri kümesini % 30 tests ve % 70 eğitimine böldü. Suç niteliğini tahmin etmek için Decision Tree ve Naive Bayesian olmak üzere iki sınıflandırma algoritması arasında bir karşılaştırma yaptık: Düşük, Orta, Yüksek. Karar Ağacı için, Doğruluk, Hassasiyet ve Geri Çağırma vardır. % 69, % 40 ve % 47. Naive Bayesian için Doğruluk, Hassasiyet ve Geri Çağırma değerleri sırasıyla % 61, % 33 ve % 19'dur. Her iki algoritma için deneysel sonuç, Karar Ağacı'nın WEKA kullanarak suç veri seti için Naive Bayesian'tan daha iyi performans gösterdiğini göstermektedir. Bu deney, 10 kat çapraz doğrulama kullanılarak gerçekleştirildi. Suç ve teröre karşı savaşla etkin bir şekilde mücadele etmek için Karar Ağacı gibi makine öğrenme algoritmalarını kullanarak kolluk kuvvetlerinin büyük avantaj sağlayabileceği açıktır. Gelecekteki araştırmalar için, suç veri kümesine diğer sınıflandırma algoritmalarını daha fazla uygulamak ve tahmin performanslarını değerlendirmek için bir plan vardır. Gelecekteki çalışmalar için başka bir yön, özellik seçimi için diğer teknikleri kullanmak ve farklı algoritmaların tahmin performansı üzerindeki etkiyi incelemektir.

Kümeleme tekniklerini kullanarak suç kalıplarını tanımlamak için veri madenciliğinin kullanılmasına baktık. Buradaki katkımız, suç kalıp tespitini makine öğrenimi görevi olarak formüle etmek ve oraya suç çözmede polis dedektiflerini desteklemek için veri madenciliğini kullanmaktı. Önemli özellikleri belirledik ; uzman tabanlı yarı denetimli öğrenme yöntemi kullanarak ve önemli nitelikleri ağırlıklandırmak için bir program geliştirdi. Modelleme tekniğimiz, suç dedektifleri için işi kolaylaştıran çok sayıda suçtan suç modellerini tanımlamayı başardı. Çalışmamızın bazı sınırlamaları, suç analizinin yerine sadece dedektife yardımcı olabileceğini de içermektedir. Ayrıca veri madenciliği, hatalı, eksik bilgiye sahip, veri girişi hatasına eğilimli olabilir. Olabilecek giriş verilerinin kalitesine duyarlıdır. Ayrıca gerçek verileri veri madenciliği özelliklerine eşlemek her zaman kolay bir iş değildir ve genellikle yetenekli veri

madencisi ve suç veri analisti gerektirir. İlk aşamalarda dedektifle yakın bir şekilde çalışmaları gerekir. Bu çalışmanın gelecekteki bir uzantısı olarak, polis kaynaklarının en etkin şekilde kullanılmasına izin vermek için, belirli bir zaman aralığı için polisin muhtemel suç yerlerine konuşlandırılmasına yardımcı olacak suç sıcak noktalarını tahmin etmek için modeller oluşturacağız. Ayrıca, suçluları, şüphelileri, çeteleri birbirine bağlamak ve ilişkilerini incelemek için sosyal bağlantı ağları geliştirmeyi planlıyoruz. Buna ek olarak, suç tespitine yardımcı olmak için bölgesel verilere gerek duyulmaktadır (kolluk veritabanlarında şüpheli açıklama arama, farklı şehirlerden trafik ihlali veritabanlarına vb. gibi).

KAYNAKLAR

Almanie, T., Mirza, R., & Lor, E. (2015). Crime prediction based on crime types and using spatial and temporal criminal hotspots. arXiv preprint arXiv:1508.02050.

Asmai, S. A., Roslin, N. I. A., Abdullah, R. W., & Ahmad, S. (2014). Predictive crime mapping model using association rule mining for crime analysis. Age, 12(21), 1.

Atkin, H. (2011). Criminal Intelligence: Manual for Analysts.

Austin, R., Cooper, G., Gagnon, D., Hodges, J., Martensen, K., & O'Neal, M. (1973). Police crime analysis unit hand-book. National Institute of Law Enforcement and Criminal Justice, Washington, DC.

Baboo, S. S. (2011). An enhanced algorithm to predict a future crime using data mining. International Journal of Computer Applications, 975, 8887.

Boba, R. (2001). Introductory guide to crime analysis and mapping. Community Oriented Policing Services. USA.

Buck, G. A. (1973). Police crime analysis unit handbook. US Department of Justice, Law Enforcement Assistance Administration, National Institute of Law Enforcement and Criminal Justice.

Chen, H., Chung, W., Xu, J. J., Wang, G., Qin, Y., & Chau, M. (2004). Crime data mining: a general framework and some examples. computer, 37(4), 50-56.

David, H., & Suruliandi, A. (2017). SURVEY ON CRIME ANALYSIS AND PREDICTION USING DATA MINING TECHNIQUES. ICTACT journal on soft computing, 7(3).

Douglas, J. E., Ressler, R. K., Burgess, A. W., & Hartman, C. R. (1986). Criminal profiling from crime scene analysis. Behavioral Sciences & the Law, 4(4), 401-421.

Estivill-Castro, V., & Lee, I. (2001, September). Data mining techniques for autonomous exploration of large volumes of geo-referenced crime data. In Proc. of the 6th International Conference on Geocomputation (pp. 24-26).

Goldberg, C., Champagne, D., & Singleton, H. V. (2007). Final report: Law enforcement and criminal justice under public law 280. US Department of Justice, Washington, DC.

Gottlieb, S., & Arenberg, S. I. (1991). Crime Analysis: From Concept to Reality. Office of Criminal Justice Planning.Gottschalk, P. (2016). Investigation and prevention of financial crime: Knowledge management, intelligence strategy and executive leadership. CRC Press.

Grossrieder, L., Albertetti, F., Stoffel, K., & Ribaux, O. (2013). Des données aux connaissances, un chemin difficile: réflexion sur la place du data mining en analyse criminelle. Revue internationale de criminologie et de police technique et scientifique, 66(1), 99-116.

Hassani, H., Huang, X., Silva, E. S., & Ghodsi, M. (2016). A review of data mining applications in crime. Statistical Analysis and Data Mining: The ASA Data Science Journal, 9(3), 139-154.

Hosseinkhani, J., Koochakzaei, M., Keikhaee, S., & Naniz, J. H. (2014). Detecting suspicion information on the Web using crime data mining techniques. International Journal of Advanced Computer Science and Information Technology, 3(1), 32-41.

Ivan, N., Ahishakiye, E., Omulo, E. O., & Taremwa, D. (2017). Crime Prediction Using Decision Tree (J48) Classification Algorithm.

Johnsen, J. W. (2016). Algorithms and Methods for Organised Cybercrime Analysis (Master's thesis).

Jung, J., & Lee, J. (2017). Contemporary financial crime. Journal of Public Administration and Governance, 7(2), 88-97.

Kalmegh, S. (2015). Analysis of weka data mining algorithm reptree, simple cart and randomtree for classification of indian news. International Journal of Innovative Science, Engineering & Technology, 2(2), 438-446.

Kaur, N. (2016). Data Mining Techniques used in Crime Analysis:-A Review. International Research Journal of Engineering and Technology (IRJET) Volume, 3.

Kiani, R., Mahdavi, S., & Keshavarzi, A. (2015). Analysis and prediction of crimes by clustering and classification. International Journal of Advanced Research in Artificial Intelligence, 4(8), 11-17.

Krishnamurthy, R., & Kumar, J. S. (2012). Survey of data mining techniques on crime data analysis. International Journal of Data Mining Techniques and Applications, 1(1), 47-49.

Local Initiatives Support Corporation (LISC), & United States of America. (2017). Crime Analysis for Non-Criminal Justice Researchers: How to Find and Utilize Data to Inform Community Public Safety Initiatives.

Qayyum, S. (2018). A Survey of Data Mining Techniques for Crime Detection. University of Sindh Journal of Information and Communication Technology, 2(1), 1-6.

ÖZGEÇMIŞ

Adı Soyadı : Sientchanwa SORO

Doğum Yeri : Fildisi Sahili

Doğum Tahiri : 03.04.1987

Medeni Hali : Nissanli

Yabancı Dili : İngilizce, Fransızca, Senoufo, Bambara

Eğitim Durumu

Lise : Sikasso LAGS lisesi (2008)

Lisans : Technolab-ista Enstitüsü

BT yönetimi Anabilim Dalı (Eylül 2015 - Temmuz 2011)

Yüksek Lisans : Technolab-ista Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı (Eylül 2008 - Temmuz 2012)

Yüksek Lisans : Ankara Üniversitesi Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı (Eylül 2015 - Ağustos 2020)

Çalıştığı Kurumlar

Veritabanları Yöneticisi, City Messenger Mali, 2011 – 2013

Stok Yöneticisi, Alliance Ivory Coast, 2013 – 2015

Programci, Serbest çalışan, 2017-2019

Dış Ticaret, LC-SHOPPING, 2019-2020