





BYM308 YAPAY ZEKAYA GİRİŞ
DÖNEM İÇİ PROJE RAPORU
UyuşturConnect

GRUP NO	4
GRUP ÜYELERİ	1)Vesile HAN 2)Tuğba USLU 3)Mücahit BAYRAM

İÇİNDEKİLER

1. TAKIM ŞEMASI.....	4-5
2. PROBLEME EN YAKIN ÇÖZÜM SUNAN ULUSLARARASI MAKALELERİN ÖZETİ.....	5-7
3. KULLANILMASI PLANLANAN ALGORİTMALAR.....	7-9
4. KULLANILMASI PLANLANAN DONANIMLAR VE ÖZELLİKLERİ.....	9
5. ÖZGÜNLÜK.....	10
6. PROJE PLANI VE TAKVİMİ.....	11
7. REFERANSLAR.....	12

1. TAKIM ŞEMASI

Vesile HAN

Yazılım Geliştirici

İstanbul Sağlık ve Teknoloji Üniversitesi, Bilgisayar Mühendisliği
3. Sınıf



Lise yıllarından itibaren yazılıma ilgi duymuş ve sürekli bu alana yönelik eğitimler alarak bu alanda uzmanlaşmayı hedeflemiştir. Üniversitenin ilk yıllarından itibaren yapay zeka alanına ilgi duymuş ve bu alanda kendini geliştirmek istemiştir. Bu süre zarfında, makine öğrenmesi ve derin öğrenme gibi çeşitli projeler üzerinde çalışmıştır. Özellikle UyuşturConnect projesinde, daha önce kazandığı deneyimleri kullanarak makale ve veri bulma süreçlerinde yazılım geliştirme aşamasında aktif rol almıştır. Bu süreçte, bilimsel araştırma gereksinimlerini karşılamak için farklı veri tabanlarına erişim sağlamış ve bu verileri analiz etmek için yazılım araçları geliştirmiştir. Ayrıca, proje sürecinde ekip içinde iletişim ve işbirliği becerilerini geliştirerek başarılı bir şekilde projenin ilerlemesine katkıda bulunmuştur.

Tuğba USLU

Yazılım Geliştirici

İstanbul Sağlık ve Teknoloji Üniversitesi, Bilgisayar Mühendisliği
3. Sınıf



Tuğba'nın yazılıma olan ilgisi üniversitenin ilk yıllarında başladı ve o zamandan beri sürekli olarak bu alanda eğitimler alarak kendini geliştirmeye odaklandı. Özellikle üniversite de aldığı dersler doğrultusunda yapay zeka alanına ilgi duydu ve bu alanda doğal dil işleme ve görüntü işleme gibi çeşitli konularda projeler yaparak deneyim kazandı.

En önemli projelerinden biri olan UyuşturConnect projesinde, makale ve veri bulma ile ilgili araştırma yaparak daha önceki edindiği tecrübeleri kullanarak yazılım geliştirme sürecinde yer aldı. Bu proje, onun yapay zeka ve yazılım geliştirme alanındaki uzmanlığını daha da derinleştirdi ve gelecekteki projelerinde önemli bir tecrübe sağladı.

Mücahit BAYRAM
Yazılım Geliştirici

İstanbul Sağlık ve Teknoloji Üniversitesi, Bilgisayar Mühendisliği
3. Sınıf



Üniversitenin ilk yıllarından itibaren yazılım geliştirme çalışmaları hakkında deneyim sahibi olan Mücahit, 3. Sınıf zorunlu stajında yazılım mühendisliği bölüm başkanı ve akademisyenlerden oluşan bir ekiple çalışma fırsatı bulmuştur. Bu süreçte, görüntü işleme teknikleri kullanarak deney faresinin günlük hareketlerinin tespiti üzerine çalışmalar yapmıştır. Bu çalışmalar sonucunda görüntü işleme, yapay zekâ ve makine öğrenmesi gibi konularda deneyim sahibi olmuştur.

Yapay zekâ alanında deneyim sahibi olan Bayram, Özellikle UyuşturConnect projesinde, makale ve veri bulma süreçlerinde yazılım geliştirme aşamasında aktif rol almıştır. Bu proje kapsamında, bilimsel literatürdeki verilerin derlenmesi ve analiz edilmesinde görev almıştır. Ayrıca, projenin hedeflerine ulaşmak için ekip içinde güçlü bir iletişim ağı oluşturmuş ve işbirliğini artırarak, proje sürecinin başarılı bir şekilde ilerlemesine büyük katkı sağlamıştır.

2. PROBLEME EN YAKIN ÇÖZÜM SUNAN ULUSLARARASI MAKALELERİN ÖZETİ

a. Research on drug trafficking

Problem: Orta Doğu ve Orta Doğu'da captagon gibi amfetamin türü uyarıcıların ticareti artmakta ve methamphetamine piyasası da gelişmektedir. Bu durum, bölgedeki uyuşturucu kaçakçılığının arttığını göstermektedir.

Çözüm: UNODC, Irak ve Orta Doğu'daki uyuşturucu kaçakçılığı durumunu değerlendirmek ve ulusal, bölgesel ve uluslararası stratejilerin geliştirilmesine temel oluşturmak için bir değerlendirme gerçekleştirdi.

Veri Seti: UNODC tarafından derlenen ve ilgili hükümetlerle paylaşılan istatistiksel veriler, Irak'ta saha araştırması ile elde edilen veriler.

Sonuçlar: "Irak ve Orta Doğu'da Uyuşturucu Kaçakçılığı Dinamikleri (2019–2023): Trendler ve Yanıtlar" raporu, Orta Doğu ve Orta Doğu bölgesindeki uyuşturucu kaçakçılığı manzarasının karmaşıklığını ve dinamizmini anlamamıza yardımcı olmaktadır. Bu rapor, Irak ve bölgedeki diğer ülkelerin bu tehditle mücadele etmek için yaptığı çabaları da vurgulamaktadır.[1]

b. Detection of Illicit Drug Trafficking Events on Instagram: A Deep Multimodal Multilabel Learning Approach

Problem: Instagram gibi sosyal medya platformları, yasa dışı uyuşturucu ticaretinin önemli bir alanı haline gelmiştir. Ancak, bu tür olayları tespit etmek giderek zorlaşmaktadır.

Çözüm: Bu çalışmada, Instagram üzerindeki yasa dışı uyuşturucu kaçakçılığı olaylarını tespit etmek için derin çok modlu çok etiketli öğrenme (DMML) yaklaşımı kullanılmıştır. Model, metin ve görüntü verilerini birleştirerek yasa dışı uyuşturucuların etiketlerini tahmin etmektedir.

Veri Seti: MM-IDTE veri seti, manuel olarak etiketlenmiş çoklu ilaç etiketlerine sahip bir veri setidir.

Sonuçlar: Yapılan deneysel sonuçlar, önerilen DMML metodolojisinin, özel karakterler ve stil değişiklikleri gibi engellere rağmen IDTE'leri doğru bir şekilde tespit edebildiğini göstermektedir.[2]

c. Bağımlılık Yapıcı Madde Kullanımı Riskinin Makine Öğrenmesi Yöntemleriyle Tahmini

Problem: Bağımlılık yapıcı maddelerin kullanımıyla ilgili gelecekte hangi maddelerin kullanılma riski olduğunu tahmin etmek zor olabilir ve bu, önleyici tedbirlerin alınmasını güçleştirebilir.

Çözüm: Bu çalışma, makine öğrenme yöntemlerini kullanarak bağımlılık yapıcı maddeler kullanan bireylerin ileride hangi maddelere yönelebileceğini tahmin etmeyi amaçlamaktadır. Kullanılan yöntemler arasında KNN, Gaussian SVM ve karar ağacı bulunmaktadır.

Veri Seti: Çalışma, bağımlılık yapıcı maddelerin kullanımına ilişkin bir veri setini kullanmaktadır. Bu veri seti, bağımlılık yapıcı maddeler kullanan bireylerin geçmiş kullanım verilerini içermektedir.

Elde Edilen Sonuçlar: Yapılan uygulama sonuçları incelendiğinde, en yüksek başarı oranının %90,60 ile Gaussian SVM yöntemi elde edilmiştir. Bu sonuçlar, makine öğrenme yöntemlerinin bağımlılık yapıcı maddelerin kullanımıyla ilgili riskleri tahmin etmede etkin bir araç olabileceğini göstermektedir.[3]

d. Machine Learning Model Training and Practice: A study on Constructing a Novel Drug Detection System

Problem: Dünya genelinde uyuşturucu, AIDS ve terörizm gibi önemli kamu tehlikeleri bulunmaktadır. Bu metinde ele alınan problem, uyuşturucu kötüye kullanımının ciddi bir sosyal güvenlik sorunu olmasıdır.

Çözüm: Makine öğrenmesi teknolojisi, uyuşturucu tespitinde etkili bir araç olarak kullanılabilir. Bu çalışma, yapay zeka ve makine öğrenmesi tekniklerini temel alan uyuşturucu tespit sistemlerinin kurulmasını incelemektedir.

Veri Seti: Çalışmada kullanılan veri setleri ve bu veri setlerinin özellikleri hakkında spesifik bilgi verilmemiştir.

Sonuçlar: Makine öğrenmesi teknolojisinin uygulanmasıyla, uyuşturucu tespitinde daha etkili ve hızlı çözümler elde edilebilir. Bu çalışmanın sonuçları, gelecekte uyuşturucu kontrolü ve tespiti alanında yapılacak çalışmalara ışık tutabilir.[4]

3. KULLANILMASI PLANLANAN ALGORİTMALAR

Projede pek çok algoritma ve model kullanılarak her biri için ayrı ayrı gözlem yapılmış bu doğrultuda grafiksel ve sayısal istatistikler kayıt altına alınmıştır. Bu doğrultuda proje için en uygun modeller tespit edilmiştir. Kullanılan tüm modeller ve işlevleri aşağıdaki gibidir:

a. Concept of backpropagation

Backpropagation, "hataların geriye doğru yayılması" anlamına gelen temel bir derin öğrenme algoritmasıdır. Yapay sinir ağlarının denetimli öğrenme sürecini, gradyan inişi kullanarak kolaylaştırır. Bir yapay sinir ağı ve bir hata fonksiyonu verildiğinde, backpropagation yöntemi bu hata fonksiyonunun sinir ağındaki ağırlıklara göre gradyanını hesaplar. Bu gradyanı kullanarak ağırlıkları iteratif olarak ayarlar ve hatayı minimize ederek ağın performansını artırır. Backpropagation'ı anlamak, etkili sinir ağı tasarımı için önemlidir, özellikle de kaybolan gradyanlar veya pasif nöronlar gibi yaygın sorunlardan kaçınmada. Tensor Flow veya Keras gibi kullanıcı dostu araçların bulunmasına rağmen, backpropagation'ın mekanizmalarını kavramak, becerikli ağ tasarımı için önemlidir. Türev ve zincir kuralı gibi matematiksel kavramların ustalığı, bu algoritmayı anlamada yardımcı olur. Temelde, backpropagation yapay sinir ağlarının, sağlanan eğitim örneklerine dayanarak girişten çıkışa karmaşık eşlemeleri öğrenmesini sağlar.

b. Gradient Descent

Gradient descent, bir fonksiyonun (f) bir maliyet fonksiyonunu (cost) minimize eden parametrelerin (katsayılar) değerlerini bulmak için kullanılan bir optimizasyon algoritmasıdır. Genellikle, parametreler analitik olarak hesaplanamadığında (örneğin, doğrusal cebir kullanılarak) ve bunların bir optimizasyon algoritması tarafından aranması gerektiğinde gradient descent tercih edilir. Gradient descent, bir fonksiyonun minimumunu bulmak için birinci dereceden tekrarlanan bir

optimizasyon algoritmasıdır. Bir fonksiyonun yerel minimumunu gradient descent kullanarak bulmak için, mevcut noktadaki fonksiyonun negatif gradientine (veya yaklaşık gradientine) orantılı adımlar atılır. Gradient Descent'i açıklamak için klasik dağcılık örneğini kullanabilirim. Diyelim ki bir dağın tepesindeyiz ve dağın en düşük noktasında (yani vadi) bulunan bir göle ulaşmanız gerekiyor. Bir türlü bağlı gözleriniz var ve nereye gittiğinizi görmek için hiçbir görüş açınız yok. Peki, göle ulaşmak için hangi yaklaşımı izleyeceksiniz? En iyi yol, etrafınızda bulunan zemini kontrol etmek ve toprağın nereye doğru inişe geçtiğini gözlemlemektir. Bu, ilk adımınızı atmanız gereken yön hakkında bir fikir verir. Eğer iniş yolunu takip ederseniz, göle ulaşma olasılığınız oldukça yüksektir.

c. K-Nearest Neighbours (kNN)

kNN algoritmasının temel kavramı, bir nesnenin sınıfının, onun k en yakın komşusunun çoğunluğunun sınıfı olduğudur. Bu algoritma, mesafe tanımına oldukça duyarlıdır. bu projede, k-en Yakın Komşu (kNN) algoritmasını kullanmayı amaçladık çünkü bu algoritma, veri tabanındaki mevcut örneklerin sınıflandırılması için etkili bir yöntemdir. Temel olarak, kNN algoritması bir nesnenin sınıfını, bu nesneye en yakın olan k diğer örneğin sınıflarının çoğunluğuna dayanarak belirler. Bu, özellikle sınıflandırma problemlerinde basit ancak güçlü bir yaklaşım sağlar. Ayrıca, projemizde farklı mesafe ölçüleri kullanarak kNN algoritmasının performansını değerlendirmeyi amaçladık. Bu, özellikle uyuşturucu madde kullanım riskini tahmin etme gibi duyarlı bir konuda, en iyi sonuçları elde etmek için önemlidir. Son olarak, kNN algoritması projemiz için seçilmiş temel bir algoritma olduğu için, farklı versiyonlarını test etmek ve en iyi performansı elde etmek için çeşitli parametreleri ve yöntemleri inceledik. Bu sayede, projenin amacına en uygun kNN modelini belirlemeyi hedefledik.

d. Decision Tree(DT)

Karar ağacı yaklaşımı, bir ağaç yapısı oluşturan bir sınıflandırıcıdır. Bu ağaç yapısı, çeşitli eylem seçenekleri arasında seçim yapmak için kullanılabilir. Bu çalışmada, ikili karar ağaçları kullanılmıştır. Bir karar ağacı, düğümlerden ve yapraklardan oluşur. Her bir düğüm, bir çocuk düğümü olabilir. Bir düğümün çocuk düğümü yoksa, yaprak veya terminal düğüm olarak adlandırılır. Herhangi bir karar ağacı, hiçbir ebeveyn düğümü olmayan bir kök düğüm içerir. Her terminal olmayan düğüm kendi Boolean ifadesini hesaplar ('true' veya 'false' değerleriyle). Bu hesaplamanın sonucuna göre, belirli bir örneğe ilişkin karar, sol çocuk düğümüne ('true') veya sağ çocuk düğümüne ('false') atanır. Her yaprak (terminal düğüm), eğitim setindeki örneklerin her sınıfa ait olduğunu gösteren bir etikete sahiptir. Her sınıfın olasılığı, bu sınıftaki örneklerin sayısının, yaprağın toplam örnek sayısına oranı olarak tahmin edilir.

Karar ağacı geliştirmek için birçok yöntem vardır. Bu çalışmada, bilgi kazancı, Gini kazancı ve DKM kazancına dayalı yöntemler kullanılmıştır. Gerçek değerli girdileri karar ağaçlarında kullanmanın çeşitli yaklaşımları vardır. Yaygın olarak kullanılan bir yaklaşım, ağacın oluşturulmadan önce gerçek değerli özniteliklerin gruplandırılmasıdır. Bu çalışmada, 'anlık' grupta kullanılmıştır: her bir düğümde

her gerçel değerli özniteliğe en iyi eşik değeri aranır ve sonra bu eşik değeri bu düğümde bu özniteliği gruplamak için kullanılır. En iyi eşik değeri, kullanılan bölme kriterine bağlıdır (bilgi kazancı, Gini kazancı veya DKM kazancı). Ayrıca, kullanılan bir diğer olasılık Fisher'ın ayırmacı işlevinin her bir düğümde gerçel değerli özelliklerin en iyi lineer kombinasyonunu tanımlamak için kullanılmasıdır.

Ağacın iyileştirilmesi için budama teknikleri uygulanmıştır. Ağacın yaprağındaki belirtilen minimal örnek sayısı, düğüm bölme işlemini durdurmak için bir kriter olarak kullanılmıştır. Ağacın her yaprağı, belirtilen bir örnek sayısından az olamaz.

e. Support Vector Machine(SVM)

Destek Vektör Makinesi (SVM), uyuşturucu madde kullanımının tespiti ve risk tahmini gibi alanlarda yaygın olarak kullanılan denetimli makine öğrenimi sınıflandırma tekniğidir. SVM, veri setindeki önemli örneklerden olan destek vektörlerini seçerek ve bu destek vektörleri aracılığıyla sınıfları mümkün olduğunca geniş bir şekilde ayıran bir doğrusal fonksiyon oluşturarak çalışır. Bu nedenle, SVM, uyuşturucu madde kullanım riskini belirlemek için bir giriş vektörünü yüksek boyutlu bir uzaya eşlemeyi amaçlar.

Uygulamada, SVM, veri setini iki sınıfa bölen en uygun hiperdüzlemi bulur. Bu hiperdüzlem, veri noktalarını iki sınıf arasında en iyi şekilde ayırmayı amaçlar. SVM, en uygun hiperdüzlemi belirleyerek, bu hiperdüzlem ile en yakın veri noktası arasındaki mesafeyi, yani marjinal mesafeyi maksimize etmeye çalışır. Bu şekilde, SVM, uyuşturucu madde kullanım riskini değerlendirirken sınıfları mümkün olduğunca iyi ayırmayı ve doğru bir şekilde sınıflandırmayı hedefler.

4. KULLANILMASI PLANLANAN DONANIMLAR VE ÖZELLİKLERİ

Proje kapsamında ana bilgisayar olarak **“Monster Abra A5”** kullanılacaktır. Bu bilgisayarın özellikleri şu şekildedir: Ekran kartı olarak **“NVIDIA GeForce RTX 3050 Ti”** bulunmaktadır. İşlemci olarak **“Intel Core i7”** işlemci bulunmaktadır. Ayrıca, bilgisayar **16 GB RAM** kapasitesine ve **512 GB SSD** depolama alanına sahiptir.

Projenin yedek bilgisayarı olarak aynı özelliklere sahip bir bilgisayar mevcuttur. Bu yedek bilgisayar, aynı kapasitelerdeki ekran kartı, işlemci, RAM ve depolama alanına sahiptir.

Ayrıca, proje kapsamında **“Google Colab”** kullanılacaktır. Google Colab, TPU (Tensor Processing Unit) veya Tesla K80 GPU üzerinde derin öğrenme uygulamaları geliştirebilen ve bu özelliği kullanıcıya ücretsiz sunan bir bulut servisidir. Yapılandırma gerektirmemesinin ve kolay paylaşım imkanı sunmasının yanı sıra GPU'lara ücretsiz erişim sağlar. Sunumumuzu Colab üzerinden gerçekleştireceğiz.

5. ÖZGÜNLÜK

Uyuşturucu kullanımı, toplumların ve bireylerin sağlığı üzerinde önemli bir tehdit oluşturan yaygın bir sorundur. Bu nedenle, uyuşturucu kullanımı ve bu alandaki araştırmalar hayati bir öneme sahiptir. Sunulan bu proje, uyuşturucu kullanımını kişilik özellikleri ile tahmin etmeyi amaçlamaktadır.

Bu projenin yenilikçi tarafı, kişilik özelliklerinin uyuşturucu kullanımı ile ilişkisini derinlemesine incelemesidir. Beş Faktör Modeli (FFM) üzerine odaklanarak, Nörotisizm, Dışadönüklük, Deneyime Açıklık, Uyum ve Dikkatli olmak üzere beş temel kişilik özelliğini değerlendirecektir. Bunun yanı sıra, uyuşturucu kullanımını öngörmek için farklı sınıflandırma yöntemleri ve makine öğrenimi algoritmaları kullanılacaktır. Elde edilen veriler, kapsamlı bir şekilde incelenecek ve model geliştirme aşamasında kullanılacaktır. Katılımcıların demografik özellikleri, eğitim seviyeleri, coğrafi konumları ve kişilik profilleri gibi faktörler dikkate alınarak, uyuşturucu kullanımı ile ilişkili desenler ve eğilimler belirlenecektir. Bu analizler, uyuşturucu kullanımını tahmin etme konusunda değerli bir içgörü sağlayacaktır.

Makine öğrenimi algoritmaları, veri setindeki ilişkileri tanımlayacak ve uyuşturucu kullanımını tahmin etmek için model oluşturacaktır. Sınıflandırma yöntemleri arasında Backpropagation, Perceptron, Karar Ağaçları, Destek Vektör Makineleri (SVM) ve k-en Yakın Komşular gibi teknikler kullanılacaktır. Bu algoritmaların performansı, hassasiyet ve özgüllük gibi metriklerle değerlendirilecek ve en etkili sınıflandırıcılar belirlenecektir.

Bu çalışmanın sonuçları, uyuşturucu kullanımıyla ilişkili risk faktörlerini anlamak ve önlemek için önemli bir katkı sağlayacaktır. Ayrıca, kişilik özelliklerinin uyuşturucu kullanımını tahmin etme yeteneği, gelecekteki araştırmalara ve müdahalelere ilham kaynağı olabilir. Bu proje, toplum sağlığını korumak ve uyuşturucu kullanımını azaltmak için bilimsel ve yenilikçi bir yaklaşım sunmaktadır.

6. PROJE PLANI VE TAKVİMİ

Aşama 1: Uyuşturucu Kullanımıyla İlgili Literatür Araştırması ve Konsept Geliştirme (20 Nisan - 25 Nisan)

Uyuşturucu kullanımının yaygınlığı, etkileri ve mevcut mücadele stratejileri üzerine literatür araştırması yapıldı.

Proje hedefleri, kapsamı ve yaklaşımı belirlendi.

Aşama 2: Uyuşturucu Kullanımı Verilerinin Toplanması ve Temizlenmesi (26 Nisan - 30 Nisan)

Uygun veri kaynakları toplandı.

Toplanan veriler temizlenerek, düzenlendi ve analiz için hazırlandı.

Aşama 3: Uyuşturucu Kullanımı Verilerinin Analizi ve Modelleme (1 Mayıs - 10 Mayıs)

Toplanan veriler analiz edilerek ve uyuşturucu kullanımı eğilimleri belirlendi.

Makine öğrenimi modelleri bulundu projeye eklenerek geliştirildi.

Model performansı değerlendirilerek ve iyileştirme çalışmaları yapıldı.

Aşama 4: Proje Sonuçlarının Sunumu ve Raporlanması (11 Mayıs - 15 Mayıs)

Proje için yapılan çalışmalar, elde edilen sonuçlar ve önerilen çözümleri içerecek şekilde bir rapor hazırlandı.

Yapılacak sunum için gerekli sunum dosyaları hazırlandı.

Aşama 5: Proje Son Kontrol ve Gözden Geçirme (16 Mayıs - 17 Mayıs)

Projenin son hali kontrol edildi ve gerekli yerlerde revizeler yapıldı.

Sunum dosyaları ve proje raporu son kez gözden geçirildi.

Tüm dökümantasyon tamamlandı.

7. REFERANSLAR

- [1] <https://www.unodc.org/unodc/es/data-and-analysis/research-on-drug-trafficking.html>
- [2] <https://dergipark.org.tr/en/pub/sabited/issue/79593/1286821>
- [3] <https://wepub.org/index.php/IJCSIT/article/view/608>
- [4] https://www.researchgate.net/publication/354065809_Detection_of_Illicit_Drug_Trafficking_Events_on_Instagram_A_Deep_Multimodal_Multilabel_Learning_Approach