



Veri Madenciliği

Giriş

Hakkında

İçerik

Kapanış

Merhaba.

Emre Can Öner

-

emrecanoner@outlook.com

Veri Madenciliği

Giriş

Hakkında

İçerik

Kapanış

Social Media Sentiment Analysis

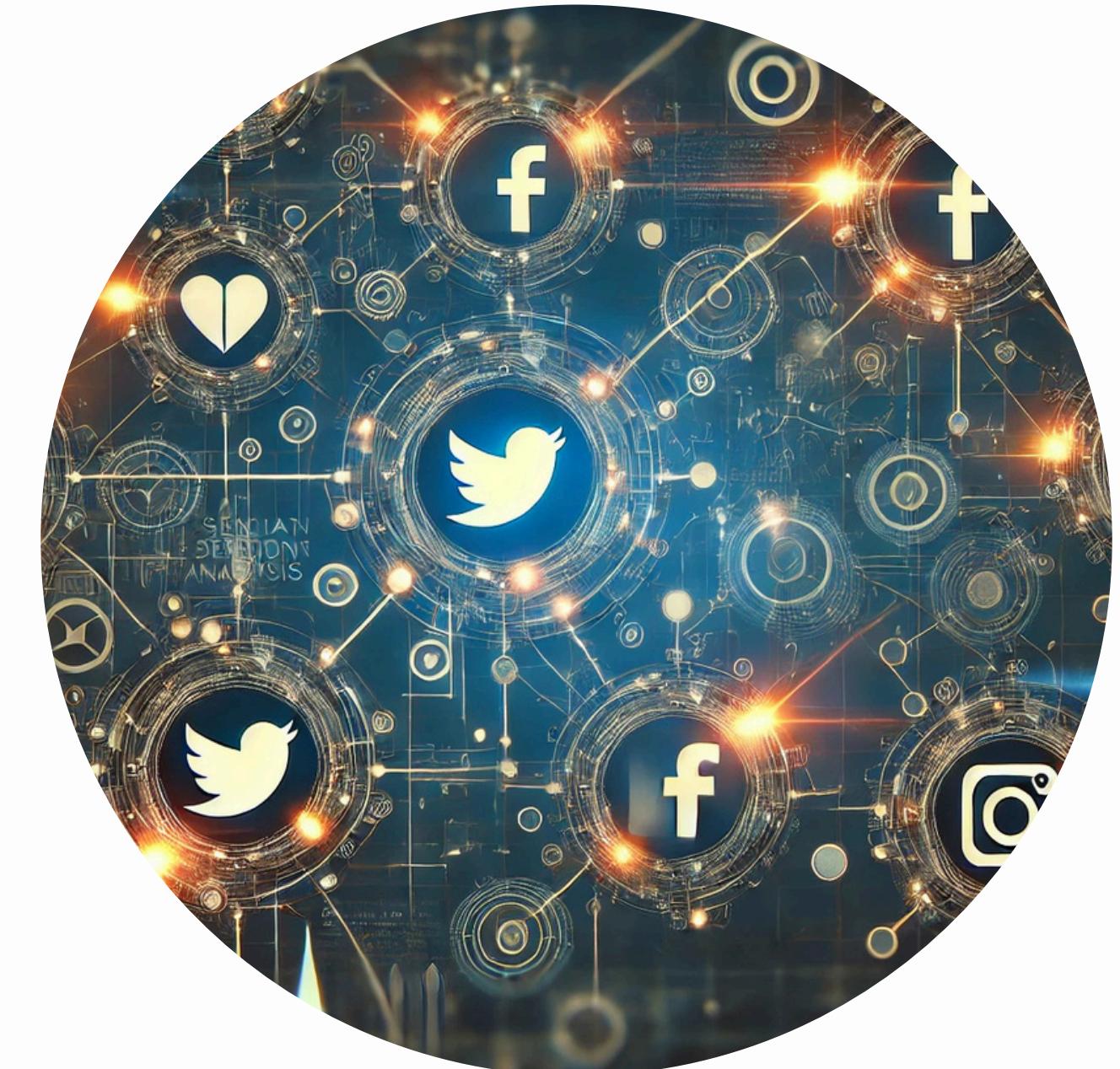


KONU

Sosyal medya duygusal analizinde transformer tabanlı modellerin kullanılması

BAĞLAM

Sosyal medya platformları, çeşitli konularla ilgili kamuoyunun duygularını yansıtan büyük miktarda veri üretir. Bu nedenle, sosyal medya duygusal analizi pazarlama, siyaset ve halk sağlığı gibi alanlarda kritik önem taşır.



Zorluklar

Sosyal medya verilerinin gayriresmi dili, kısaltmalar, emojiler ve kısa içerikleri, duygusal analizini geleneksel metin kaynaklarına kıyasla daha karmaşık hale getirir.

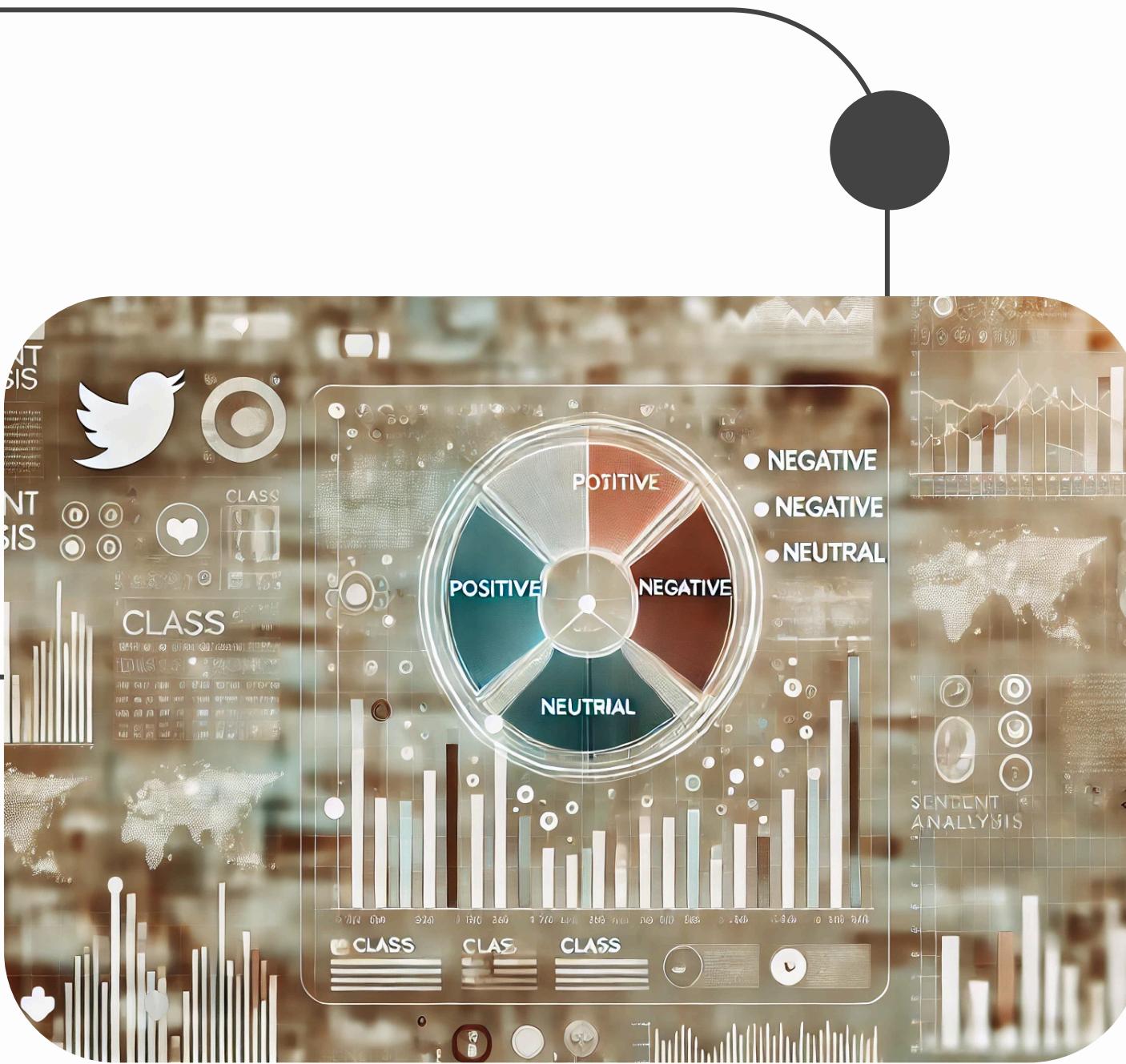
Araştırma Boşluğu

Çoğu çalışma, ya geleneksel modelleri yapılandırılmış veri setlerinde kullanır ya da transformer modellerini büyük, düzenli veri setlerine uygular. Bu proje, transformer modellerinin gayriresmi sosyal medya duygusal analizindeki etkinliğini değerlendirmeyi amaçlamaktadır.



AMAC

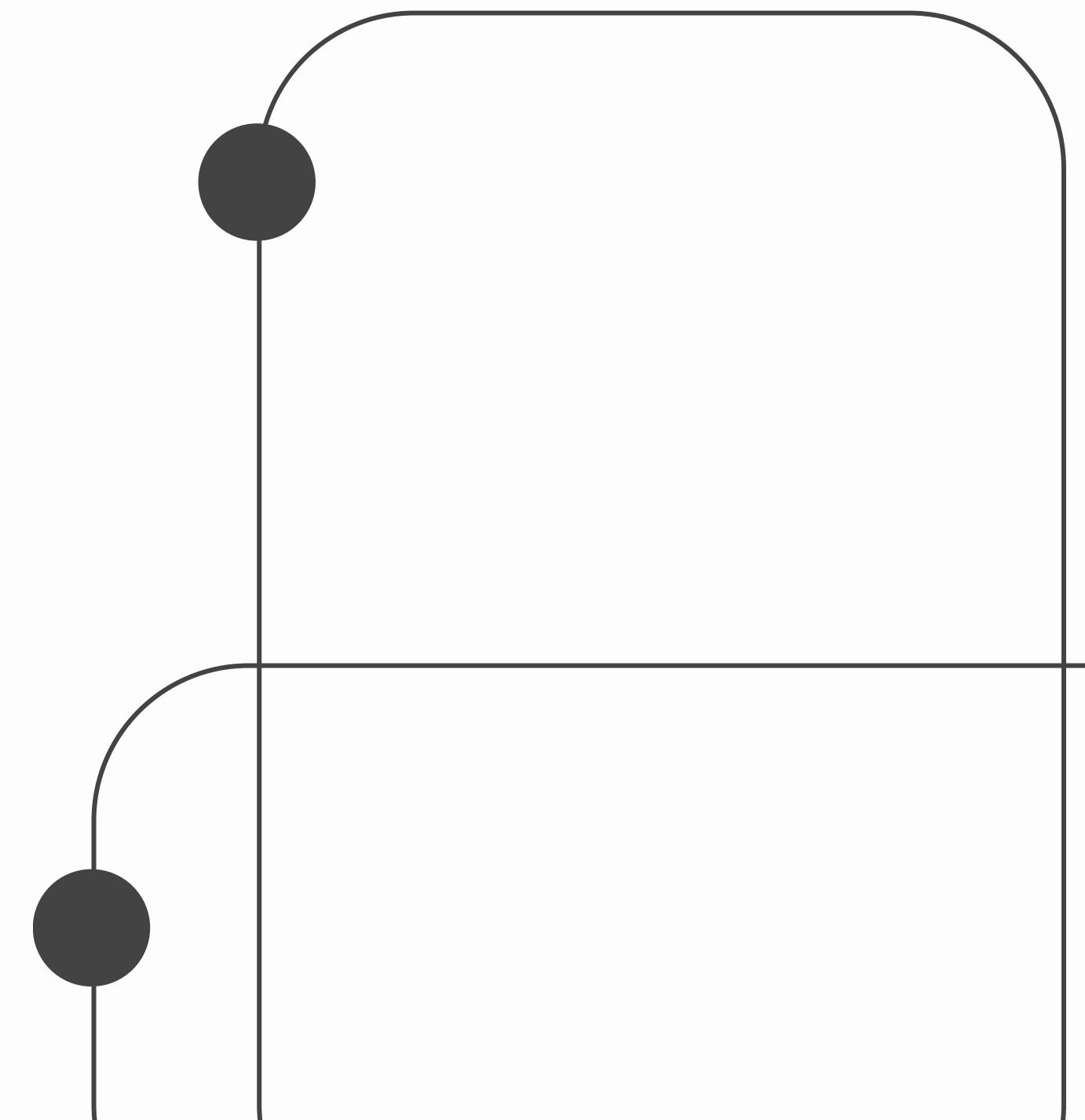
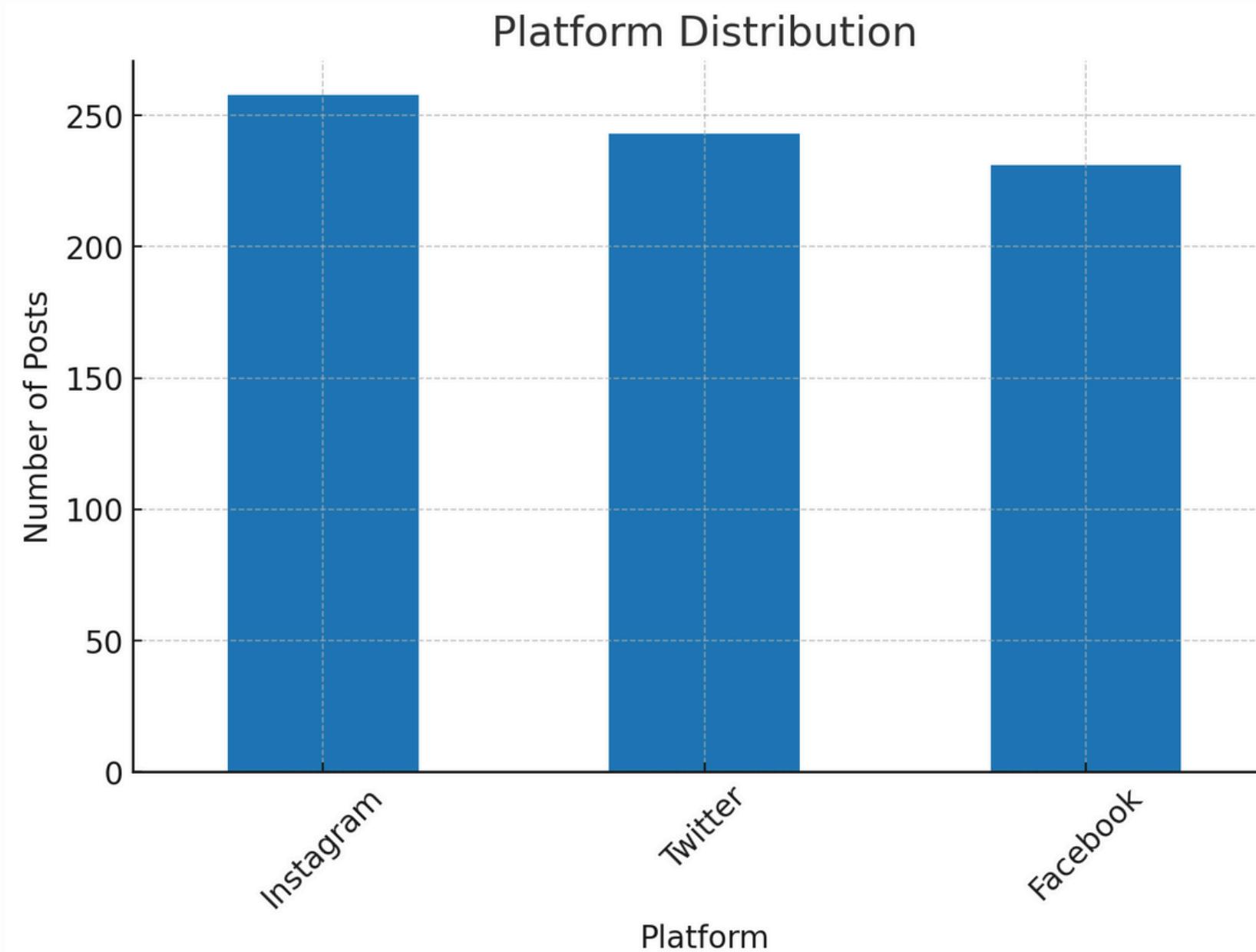
Transformer tabanlı duygusal analizi modellerinin Kaggle Sosyal Medya Duygu Analizi Veri Seti üzerindeki performansını değerlendirmek ve iyileştirmek, ayrıca optimal ön işleme tekniklerini ele almak.



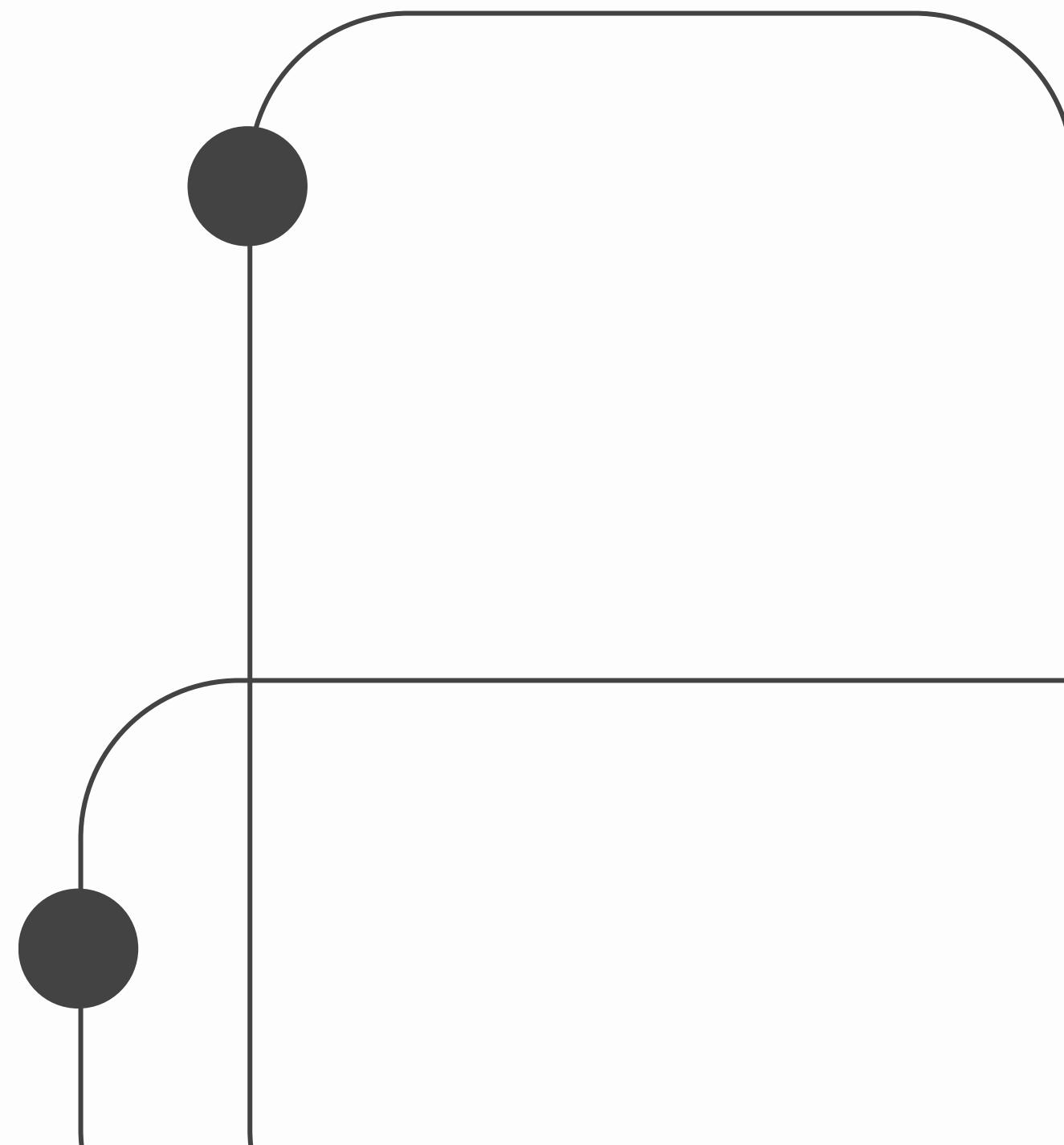
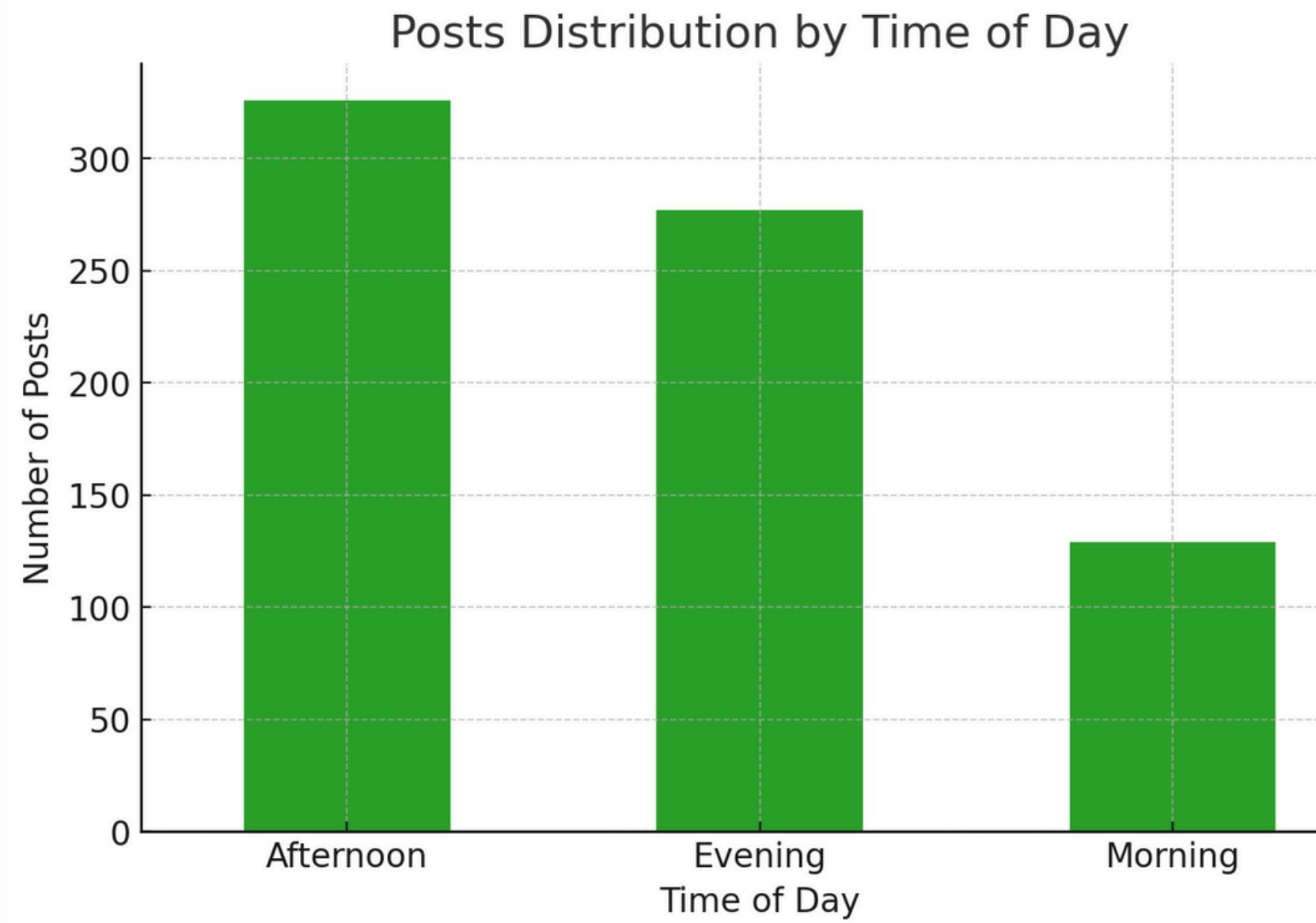
Veri Seti Analizi

Kaggle veri seti, sosyal medya gönderilerinden oluşmakta ve duygusal etiketleri içermektedir. Duygu sınıflarının (pozitif, negatif, nötr gibi) dağılımı ve sıklığı hakkında bilgi vermektedir. Bu veri seti 732 satır ve 15 sütun içermektedir.

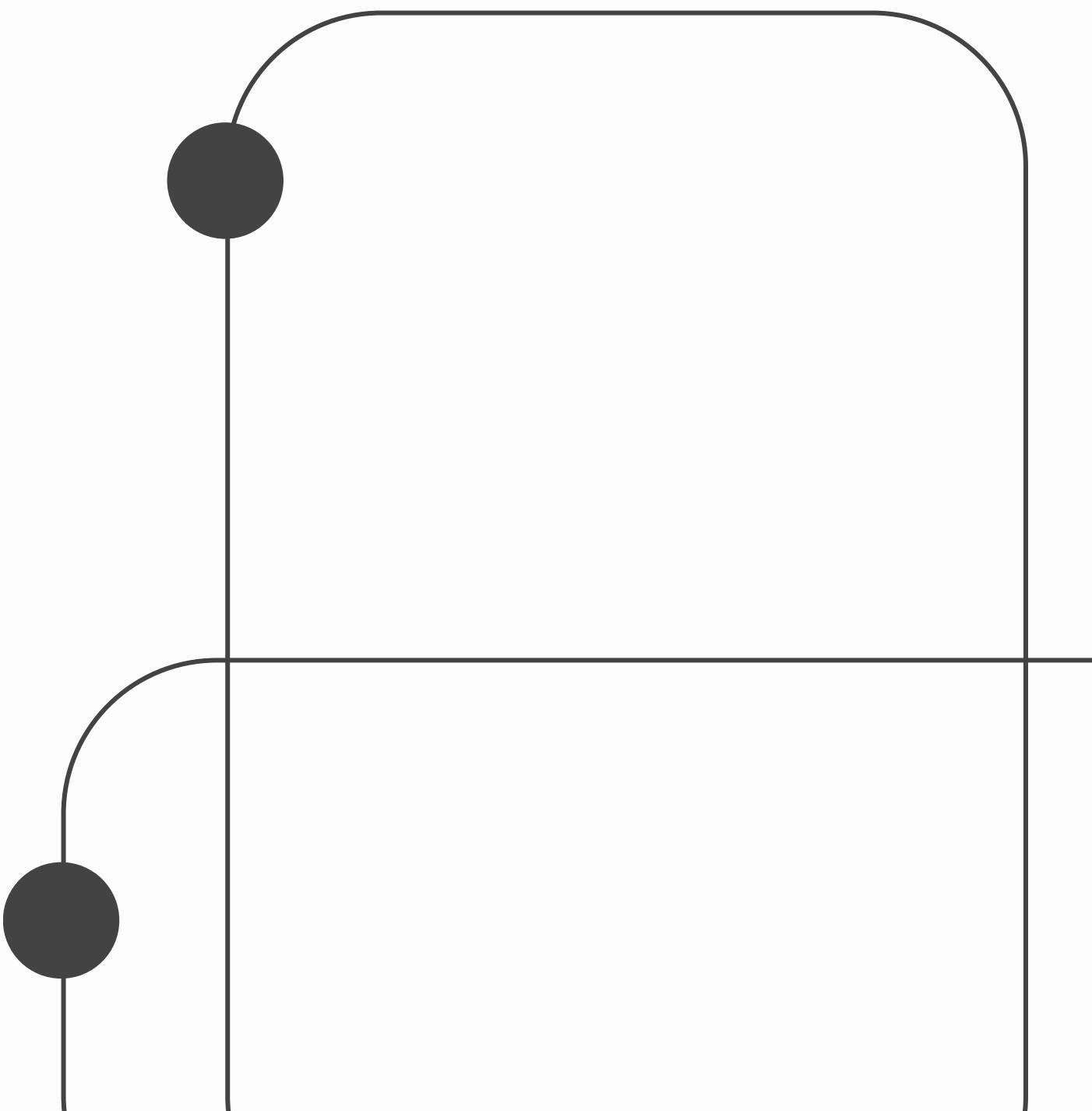
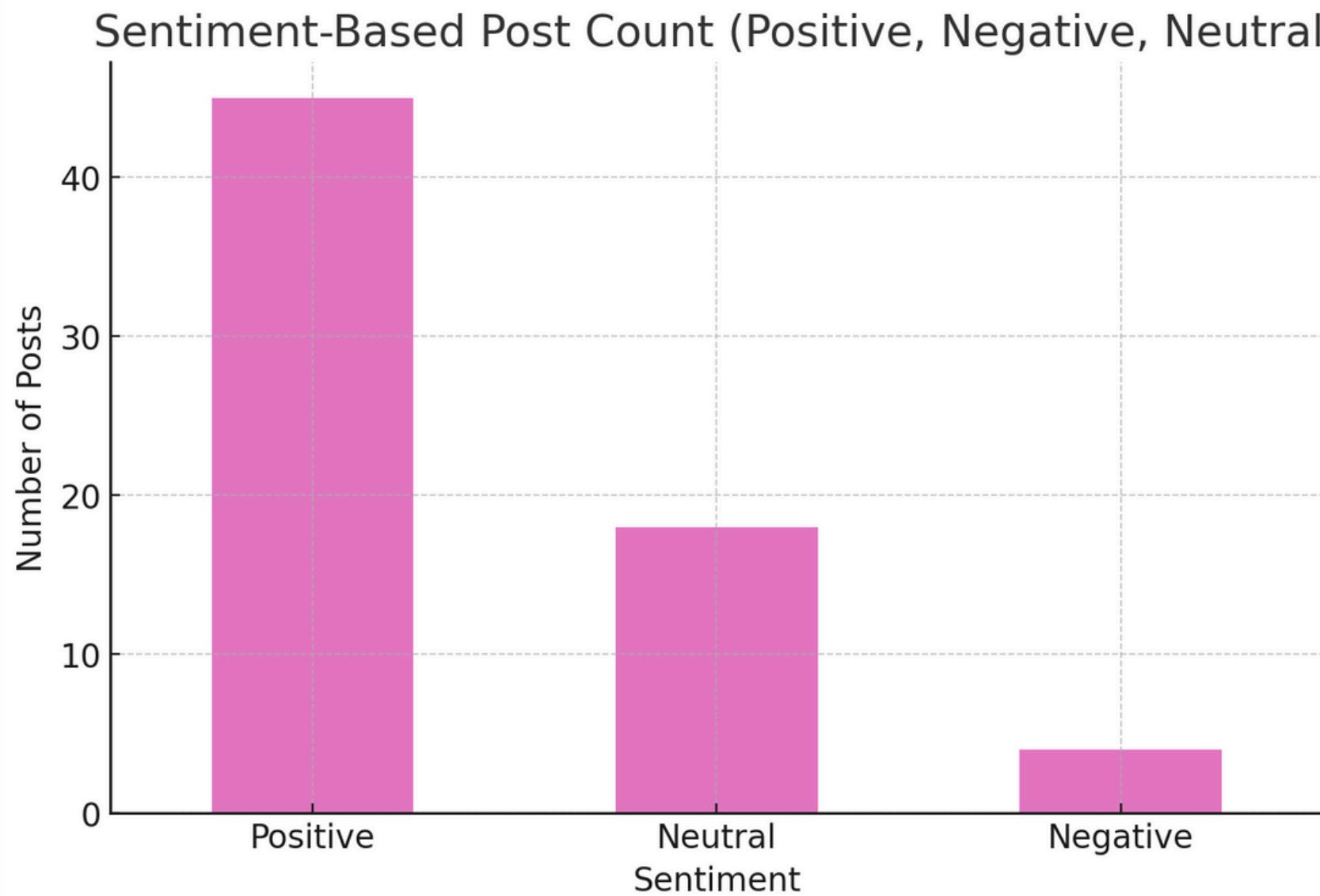
Veri Seti Analizi



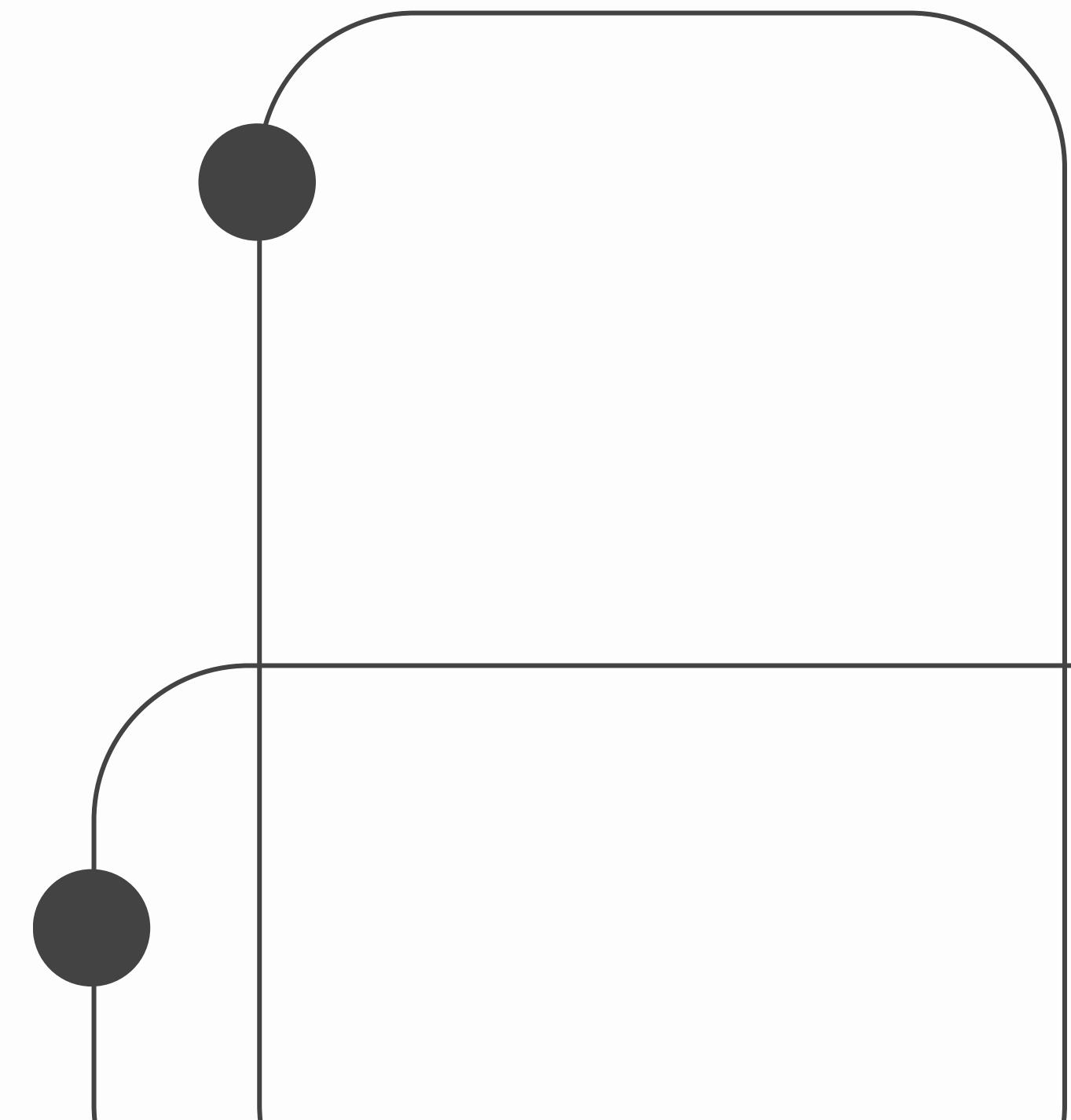
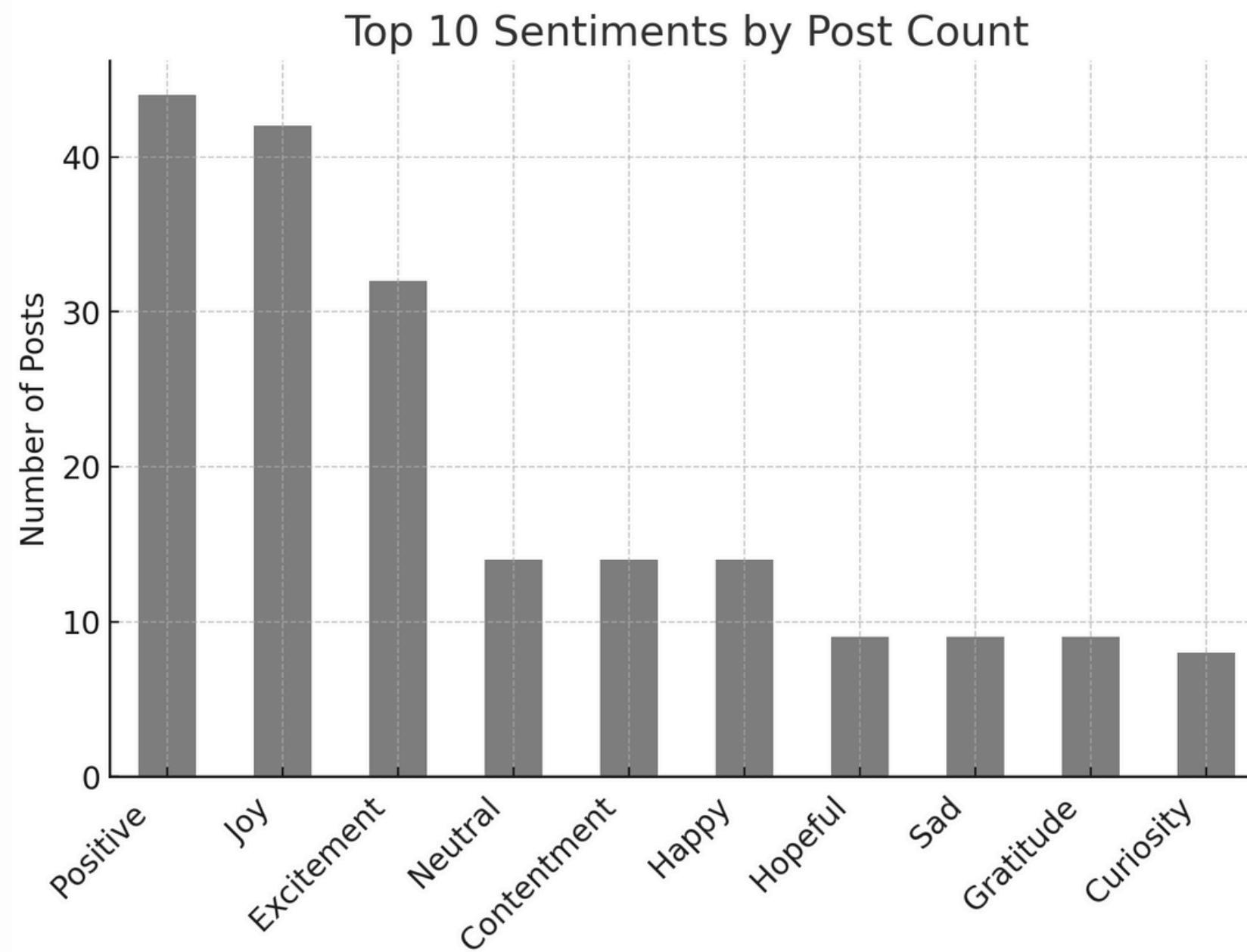
Veri Seti Analizi



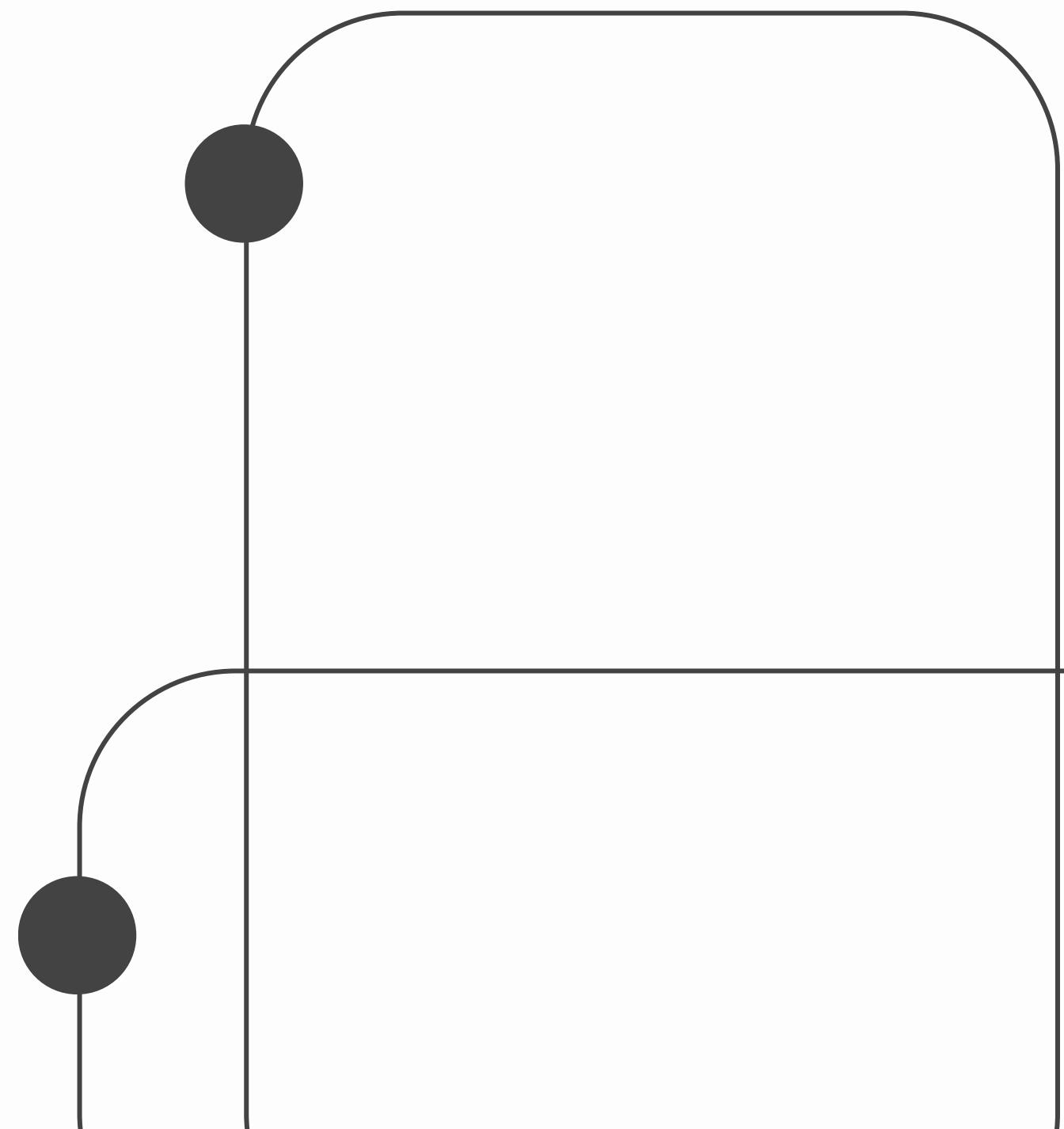
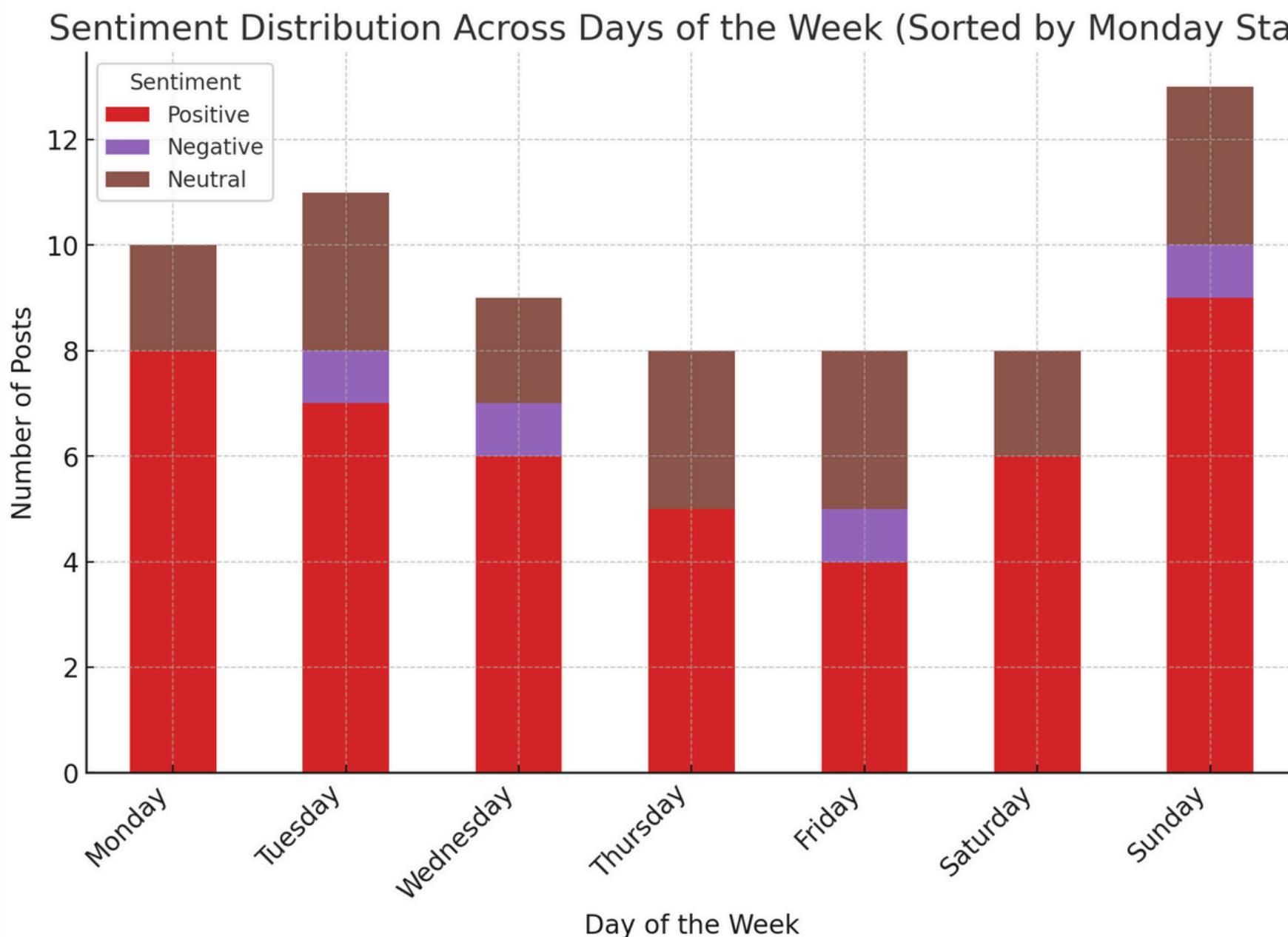
Veri Seti Analizi



Veri Seti Analizi



Veri Seti Analizi



Metodoloji

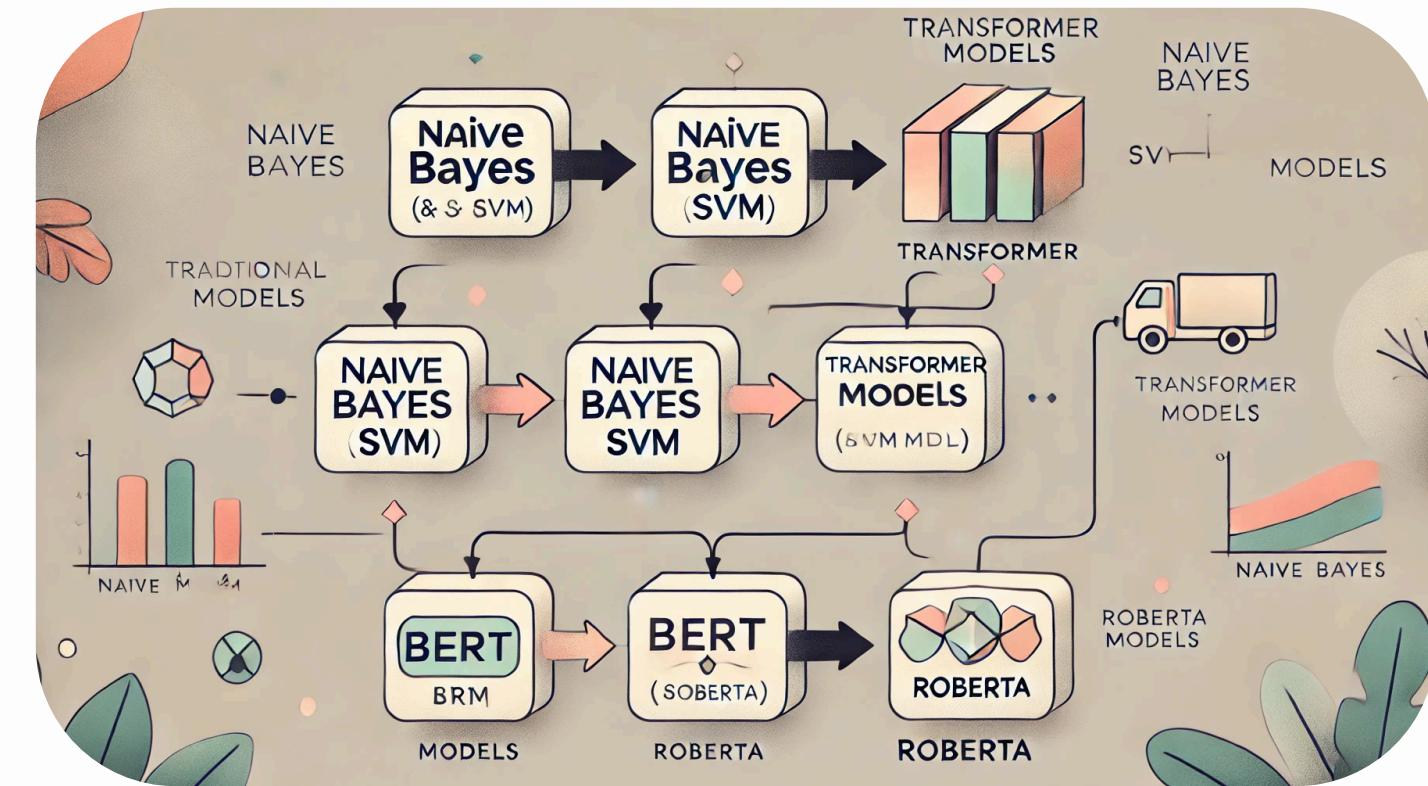
Transformer Tabanlı Modeller: Transformer modelleri, özellikle BERT ve RoBERTa, sosyal medya dilini anlamada oldukça uygundur çünkü bu modeller, kelimeler arasındaki bağlamsal ilişkileri dikkate alarak dilin anlamını derinlemesine öğrenebilirler. Sosyal medya gönderileri genellikle gayriresmi, kısaltmalar ve emojilerle dolu kısa metinlerden oluşur. Transformer modelleri, bu tür metinlerde yer alan kelimelerin farklı bağlamlarda nasıl kullanıldığını öğrenerek anlamayı derinleştirir. Ayrıca, self-attention mekanizması sayesinde metin içindeki önemli kelimelere ve ifadelere odaklanarak daha doğru duyu tahminleri yapabilirler. Bu da sosyal medya gibi değişken ve gürültülü dil yapılarında bu modelleri oldukça güçlü kılar.

Temel Modeller: Geleneksel NLP modelleri kullanılarak başlangıç performansı ölçülebilir. Bu modellerin güçlü ve zayıf yönleri, transformer modellerinin üstünlüğünü vurgulamak amacıyla değerlendirilebilir. (Naive Bayes ve SVM)

Ön İşleme Teknikleri

Tokenizasyon, metni daha küçük birimlere ayırma işlemidir ve sosyal medya verilerinde kelime varyasyonlarını ele almak için oldukça önemlidir. Transformer modelleri genellikle "alt kelime tokenizasyonu" adı verilen bir yöntem kullanır. Bu yöntemde, kelimeler anlamlı alt bileşenlere bölünür. Örneğin, "günaydınlar" kelimesi "gün", "ay", "dın", "lar" gibi alt bileşenlere ayrılarak işlenir. Bu yöntem, sosyal medya dilindeki kısaltmalar ve yaratıcı yazımların farklı varyasyonlarını daha iyi anlamaya ve işlemede transformer modellerinin etkinliğini artırmaya yardımcı olur.

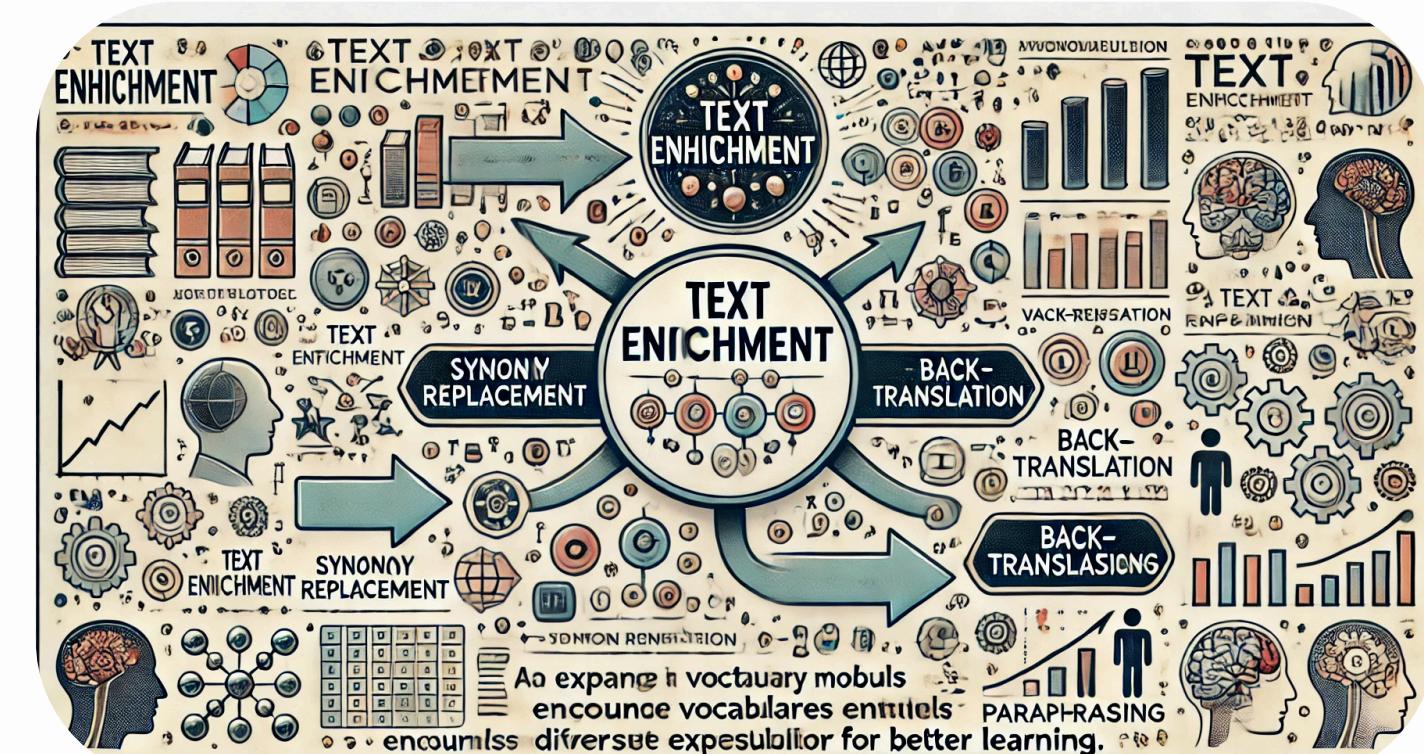
Tokenizasyon



Ön İşleme Teknikleri

Metin zenginleştirme, veri setindeki kelime dağarcığını genişletmek ve modelin daha çeşitli metinlerle karşılaşarak daha iyi öğrenmesini sağlamak için kullanılan bir tekniktir. Bu süreçte birkaç farklı yöntem kullanılabilir.

Metin Zenginleştirme



Ön İşleme Teknikleri

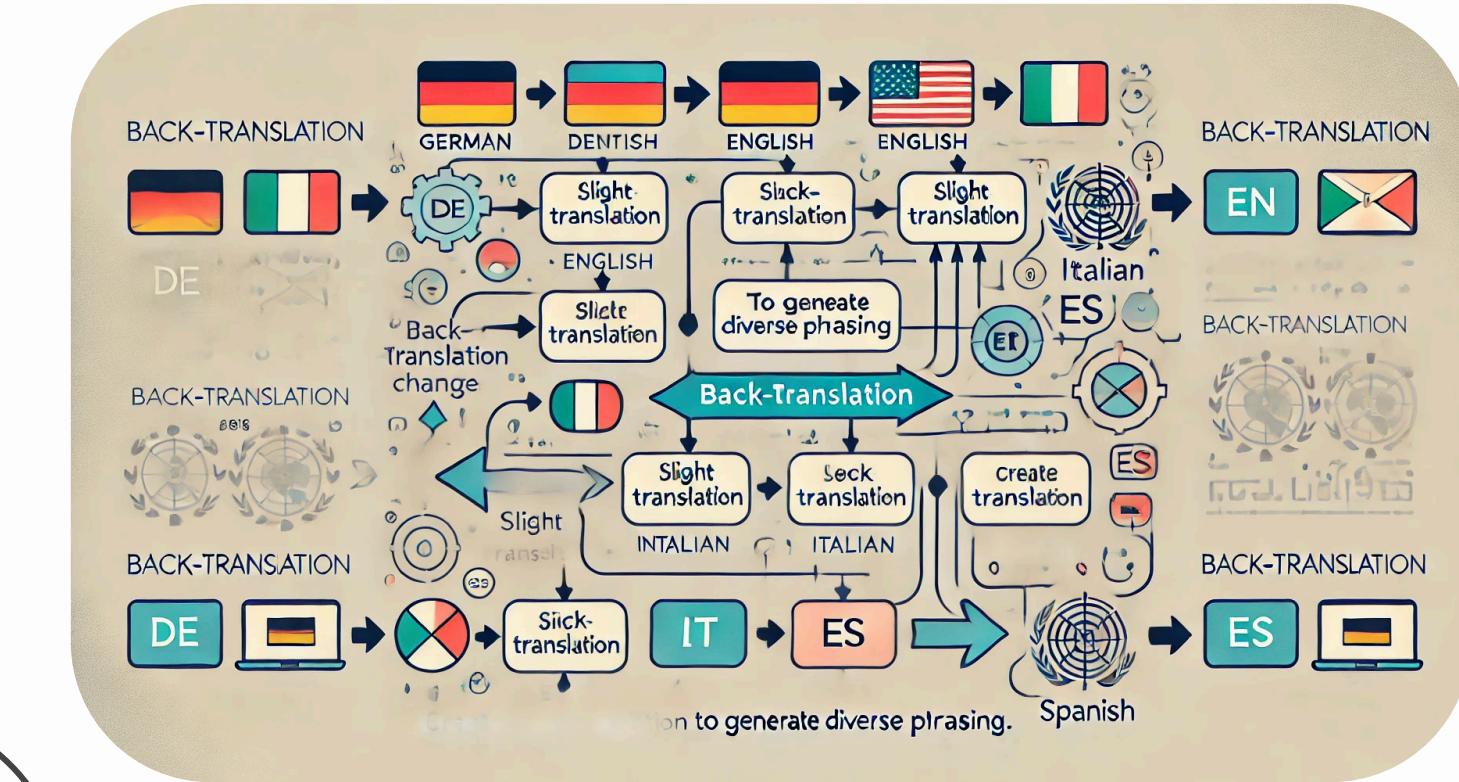
Eş anlamlı Kelime Değişimi: Mevcut metindeki bazı kelimeler, aynı anlamı taşıyan eş anlamlı kelimelerle değiştirilir. Bu yöntem, modelin farklı sözcüklerle karşılaşarak bu sözcüklerin aynı duygusal veya anlamı ifade ettiğini öğrenmesini sağlar.



Metin Zenginleştirme

Ön İşleme Teknikleri

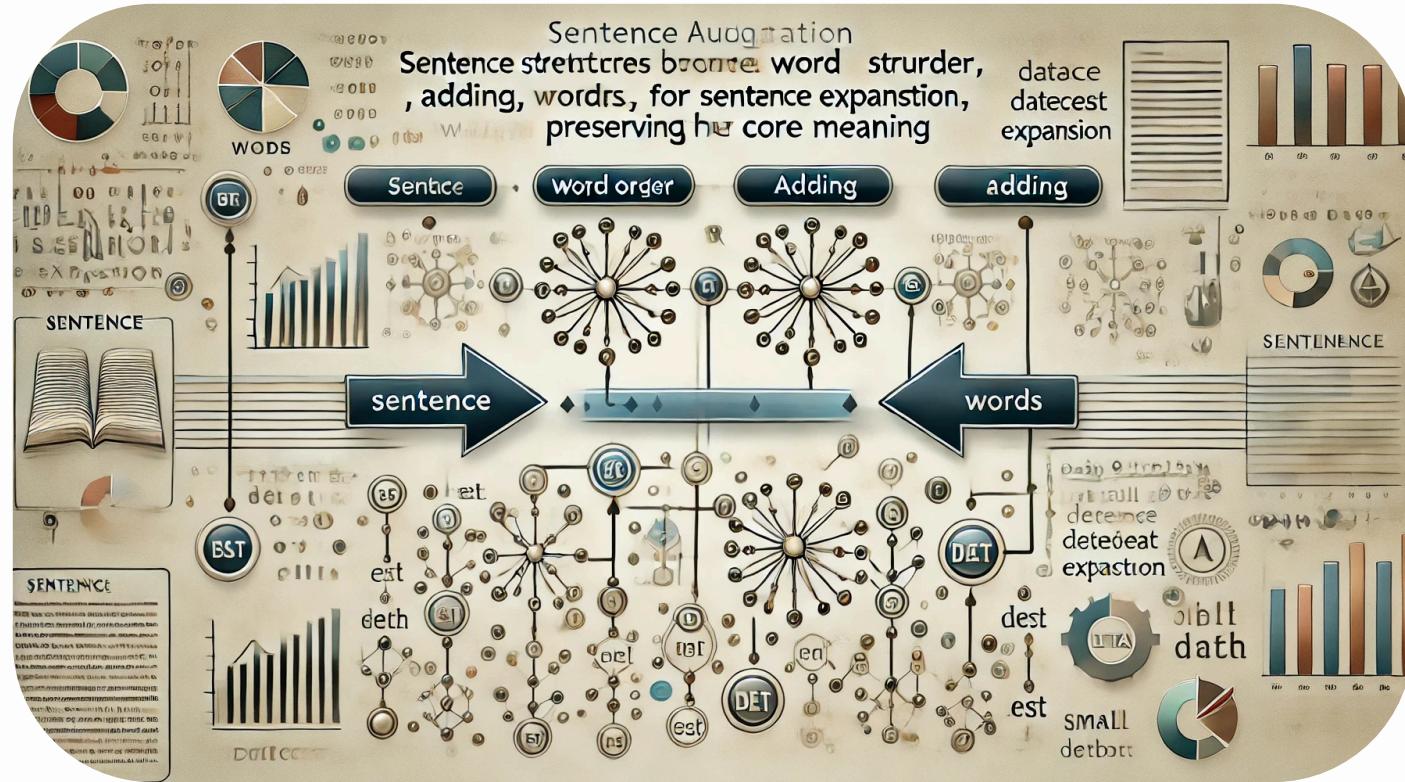
Geri Çeviri (Back-Translation): Metin, başka bir dile çevrilir ve ardından tekrar orijinal dile geri çevrilir. Bu işlem, metnin dil yapısını değiştirmeden içeriği farklı şekilde ifade etmeyi sağlayarak veri çeşitliliğini artırır.



Metin Zenginleştirme

Ön İşleme Teknikleri

Veri Çoğaltma (Data Augmentation): Kelimelerin sırasını değiştirme, rastgele kelimeler eklemeye veya çıkartma gibi yöntemlerle veri çoğaltılır. Bu, özellikle küçük veri setlerinde modelin eğitilmesi için gerekli veri miktarını artırmaya yardımcı olur.



Metin Zenginleştirme

İlgili Çalışmalar

Sosyal medya duygusal analizi üzerine yapılan birçok çalışma, farklı yaklaşımlar ve modeller kullanarak bu alandaki zorlukları ele almaktadır. Eco Green platformu gibi sürdürülebilirlik odaklı çalışmalar, CNN ve SVM kombinasyonları ile sosyal medya verilerinde duygusal analizi yaparken (Gunasekara et al., Eco Green: A Social Media Platform for Sustainable Living), Gated Fusion Semantik İlişki (GFSR) ağları kullanılarak çok modlu analizler gerçekleştirilmiştir (Zeng et al., Exploring Semantic Relations for Social Media Sentiment Analysis). Transformer tabanlı modellerin etkisi de birçok çalışmada vurgulanmış ve bu modellerin sosyal medya gibi gayriresmi dil yapısına sahip metinlerde üstün performans gösterdiği görülmüştür (Kokab & Asghar, Transformer-based deep learning models for sentiment analysis). Ayrıca, hibrit modeller (örneğin, SVM-LSA ve BERT-DCNN) duygusal sınıflandırmasında daha yüksek doğruluk sağlamak için kullanılmıştır (Kudalkar, Social Media Sentiment Analysis; Jain et al., Employing BERT-DCNN with Sentic Knowledge Base for Social Media Sentiment Analysis).

İlgili Çalışmalar

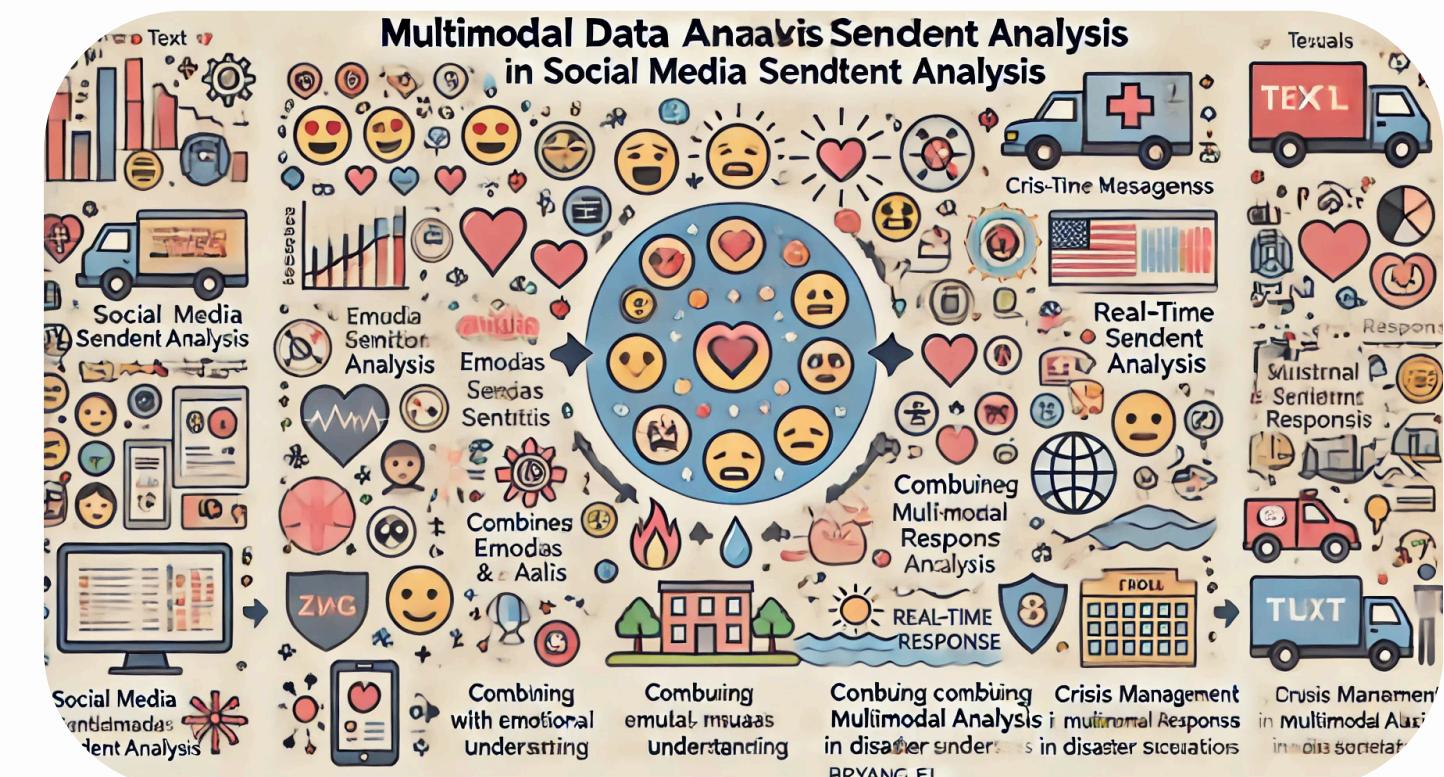
COVID-19 gibi küresel olaylar ve kriz durumlarında sosyal medya verilerinden duygusal takibi yapılarak kamu sağlığı iletişimini ve kriz yönetimi süreçlerine katkı sağlanmıştır (Rahmanti et al., Social Media Sentiment Analysis During COVID-19). Çok dilli duygusal analizi üzerine yapılan çalışmalar, sosyal medya verilerindeki dil çeşitliliğini ele alarak, birden fazla dilde duygusal sınıflandırmanın zorluklarını incelemiştir (Aguero-Torales & Salas, Deep Learning and Multilingual Sentiment Analysis). Bu kapsamda derin öğrenme modelleri, çoklu dil desteği sağlayarak daha geniş kitlelerin duygusal analizini mümkün kılmaktadır. Özellikle, dil çeşitliliği nedeniyle oluşan belirsizliklerin üstesinden gelmek için geliştirilen çapraz dil yöntemleri, sosyal medyada duygusal analizi yapılırken önemli bir rol oynamaktadır.



İlgili Çalışmalar

Ayrıca, çok modlu veri analizine yönelik yapılan araştırmalar, sosyal medyada görsel ve metinsel verilerin birleştirilmesiyle duyu analizinin doğruluğunu artırmıştır (Zeng et al., Exploring Semantic Relations for Social Media Sentiment Analysis).

Emojiler ve resimlerin metinle birlikte değerlendirilmesi, kullanıcının ifade etmek istediği duyu durumunun daha iyi anlaşılmasını sağlamaktadır. Bu tür çok modlu analizlerin kriz anlarındaki hızlı tepki gerektiren durumlarda etkili olduğu ve afet yönetimi gibi alanlarda fayda sağladığı belirtilmiştir (Bryan-Smith et al., Real-time Social Media Sentiment Analysis for Flood Impact Assessment).



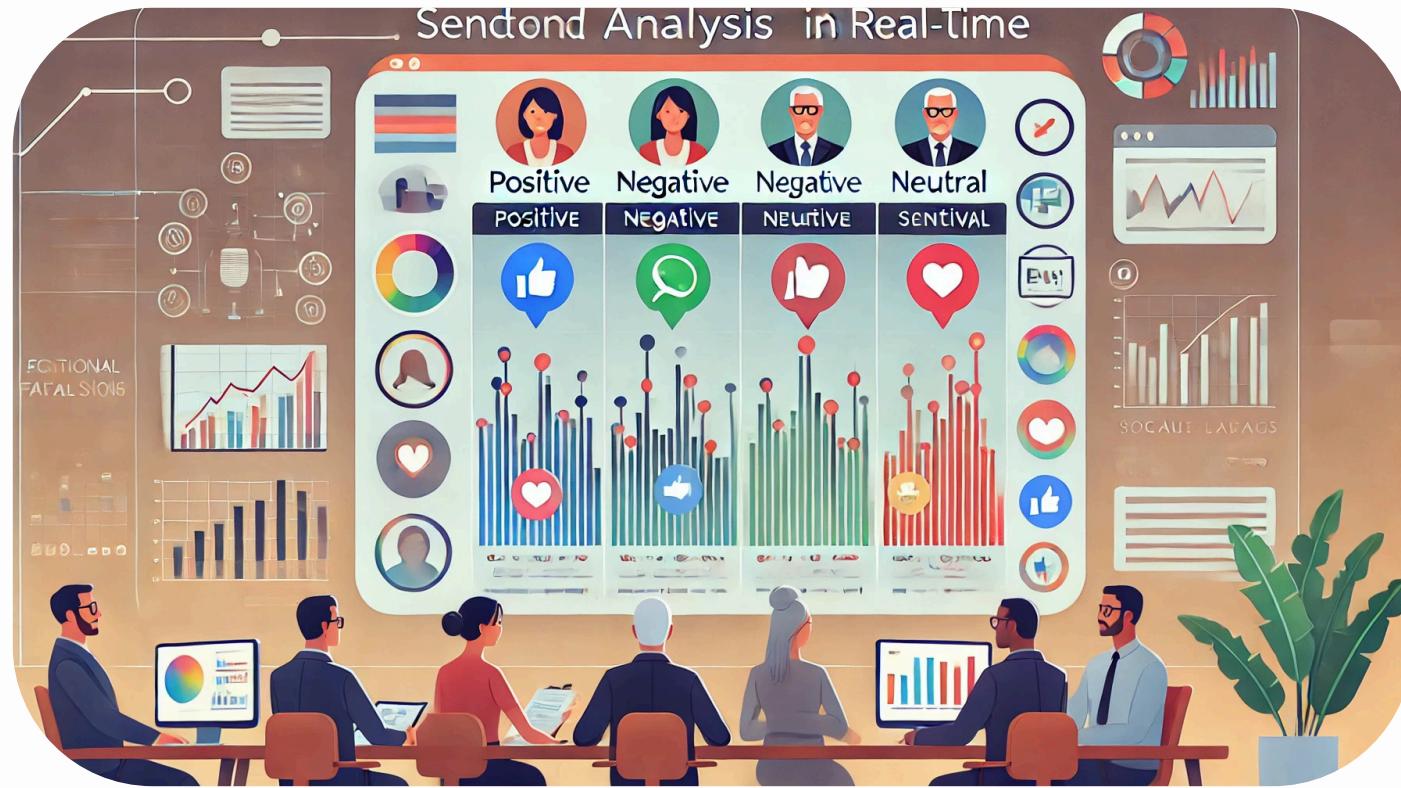
İlgili Çalışmalar

Son olarak, sosyal medya duyu analizinin ekonomik ve finansal alanlarda da kullanıldığı görülmüştür. Özellikle, Twitter gibi platformlardan elde edilen duyu verileri, kripto para piyasasındaki fiyat hareketlerini tahmin etmede kullanılmıştır (Bhatt et al., Sentiment-Driven Cryptocurrency Price Prediction). Bu çalışmalar, duyu analizinin sadece sosyal etkileşimlerde değil, aynı zamanda ekonomik göstergeler üzerinde de güçlü bir etkiye sahip olabileceğini göstermektedir.

Bu çalışmalar, sosyal medya duyu analizinde kullanılan farklı yaklaşımların geniş bir perspektifini sunmakta ve bu çalışmanın geliştirilmesinde önemli bir referans noktası oluşturmaktadır.

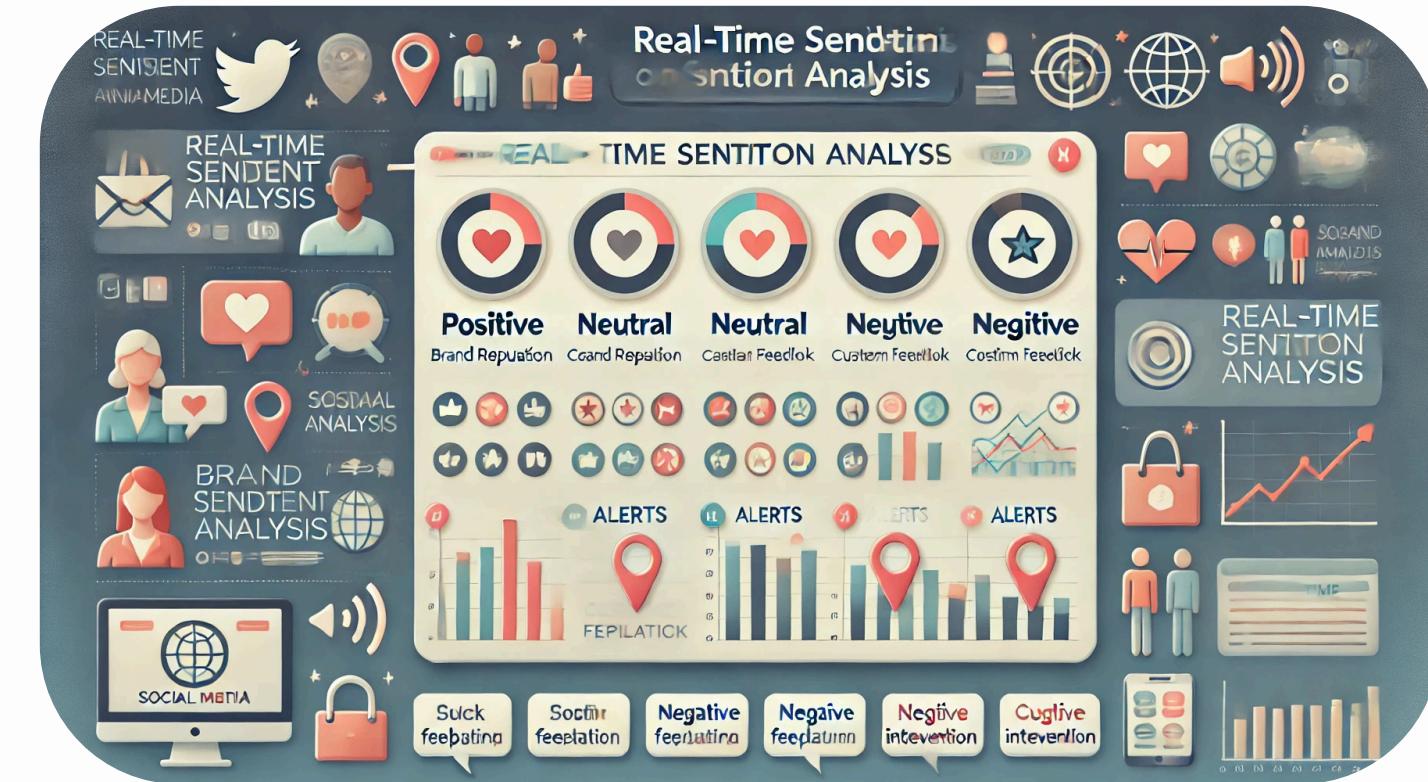
Gerçek Dünya Uygulamaları

Kamuoyu Takibi: Seçimler, etkinlikler veya ürün lansmanları gibi önemli olaylar sırasında kamuoyunun duyu durumunu gerçek zamanlı olarak izlemek ve analiz etmek için kullanılabilir. Bu sayede, kamuoyunun duygusal tepkileri hızlı bir şekilde anlaşılabılır ve bu tepkilere yönelik stratejik kararlar alınabilir. Ayrıca, politika yapıcılar ve pazarlamacılar, sosyal medya üzerinden halkın nabzını tutarak etkinliklere ve kampanyalara yönelik geri bildirimleri değerlendirme imkanı bulabilir. Bu tür bir analiz, kriz yönetimi ve halkla ilişkiler çalışmalarının daha etkili yürütülmesine de katkı sağlayabilir.



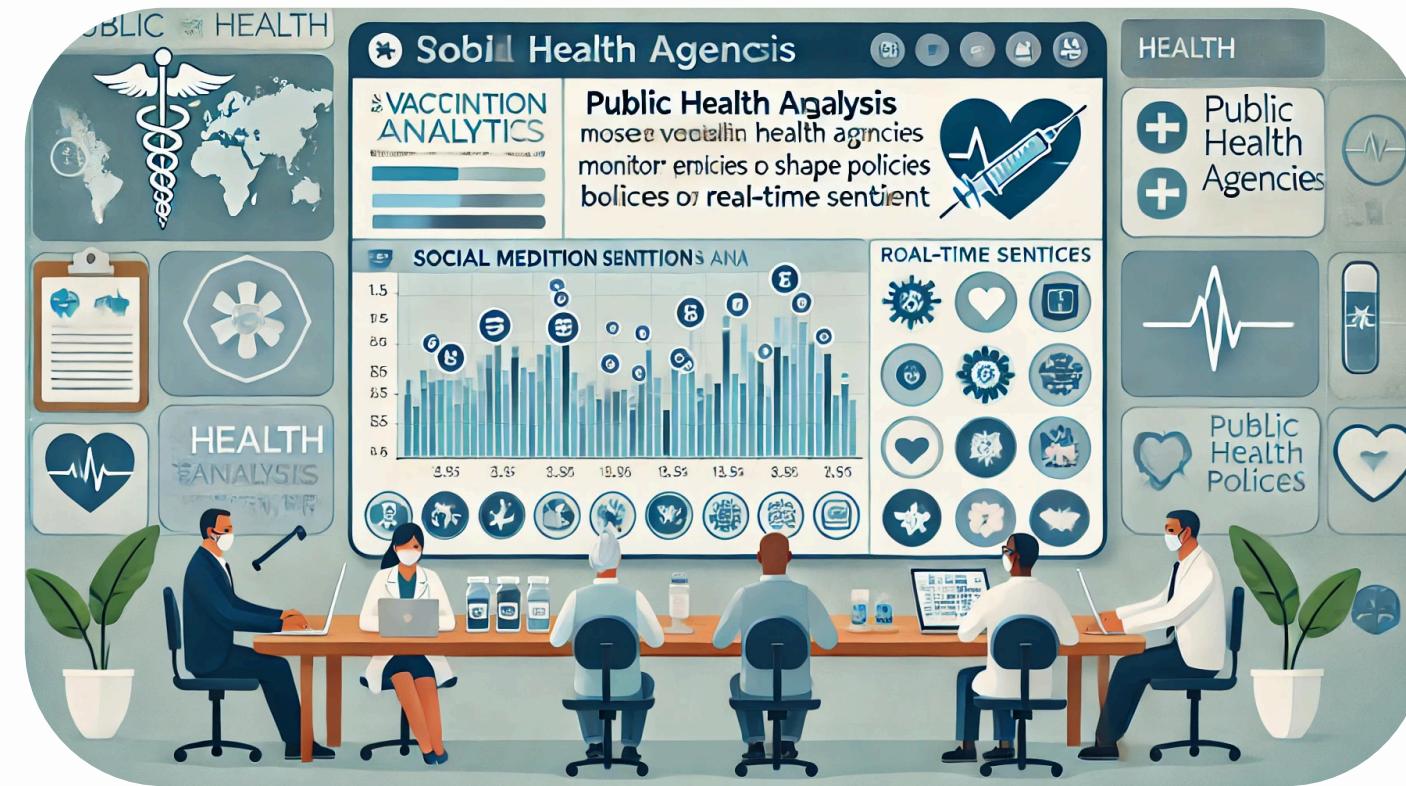
Gerçek Dünya Uygulamaları

Marka İtibar Yönetimi: Şirketlerin marka imajını ve müşteri geri bildirimlerini gerçek zamanlı olarak izleyip yönetmelerine yardımcı olabilir. Sosyal medya üzerinden elde edilen duyu analizleri, markaya yönelik olumsuz geri bildirimlerin hızlı bir şekilde tespit edilmesini ve bu durumlara anında müdahale edilmesini sağlar. Ayrıca, pazarlama kampanyalarının ve ürün lansmanlarının etkisini değerlendirerek, müşteri memnuniyetini artırmaya yönelik stratejiler geliştirilmesine katkıda bulunur. Bu, markanın itibarını koruma, müşteri bağlılığını güçlendirme ve kriz anlarında proaktif önlemler alma açısından büyük bir avantaj sağlar.



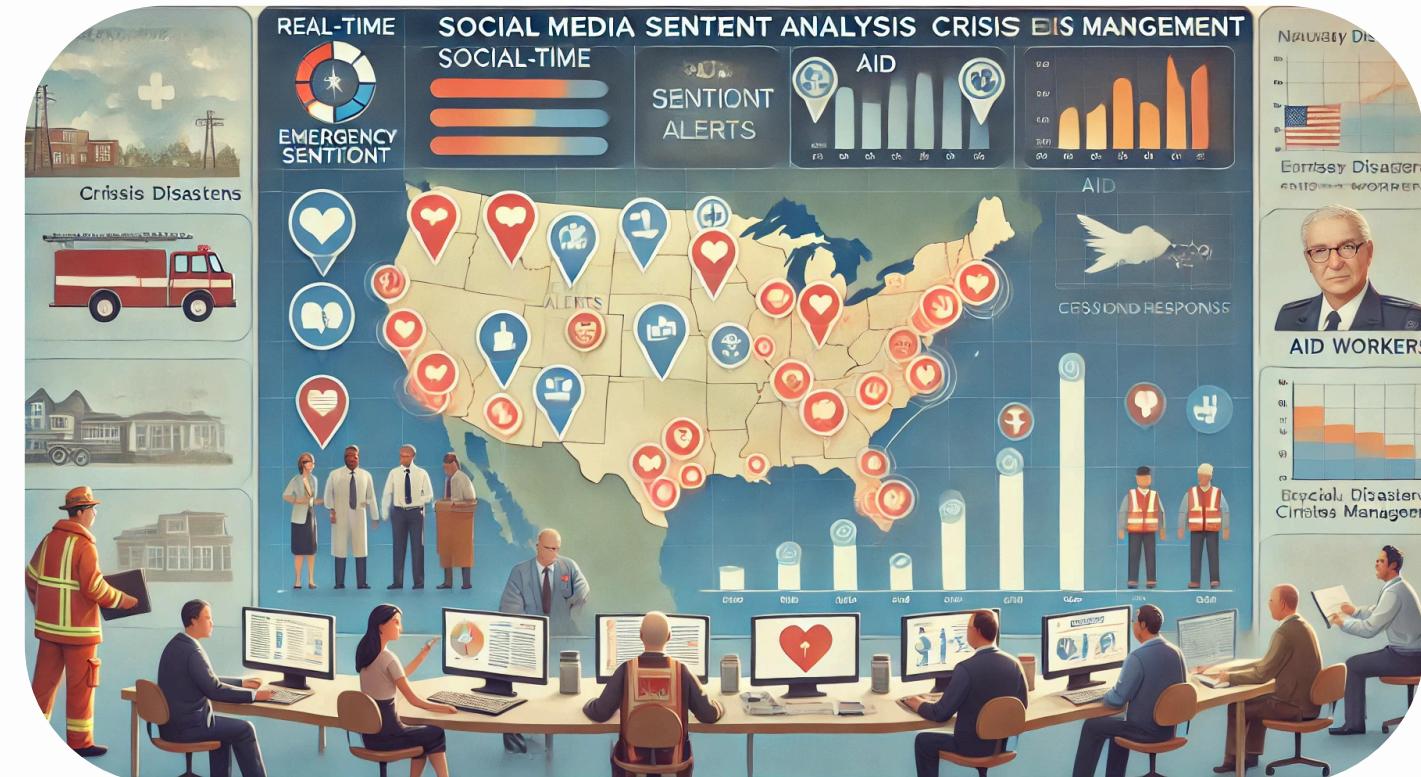
Gerçek Dünya Uygulamaları

Halk Sağlığı Takibi: Halk sağlığı ile ilgili konularda, özellikle salgın hastalıklar ve aşılama kampanyaları gibi durumlarda sosyal medyada oluşan kamuoyunun duygusal durumu izlenebilir. Bu durum, halk sağlığı politikalarının geliştirilmesine ve halkın endişelerini anlamaya yardımcı olabilir.



Gerçek Dünya Uygulamaları

Kriz Yönetimi ve Afet Müdahalesi: Doğal afetler veya kriz durumlarında, sosyal medya duyu analizi kullanılarak halkın ihtiyaçları ve duygusal durumu hızlıca analiz edilebilir. Bu, yetkililerin ve yardım kuruluşlarının, hangi bölgelerde daha fazla yardıma ihtiyaç olduğunu tespit ederek etkin bir müdahale planlamasına olanak tanır.



Gerçek Dünya Uygulamaları

Müşteri Deneyimi Optimizasyonu: Şirketler, müşteri geri bildirimlerini analiz ederek hizmetlerinde ve ürünlerinde müşteri memnuniyetini artıracak iyileştirmeler yapabilirler. Sosyal medya üzerinden elde edilen geri bildirimlerin analizi, müşteri deneyimini anlamaya ve geliştirmeye yönelik stratejilerin oluşturulmasına katkı sağlar.



Veri Madenciliği

Giriş

Hakkında

İçerik

Kapanış

Teşekkürler.

Emre Can Öner

-

emrecanoner@outlook.com