**PROJE RAPORU**

PROJE SAHİBİ: Gamze Aslan- 210541001

PROBLEM TANIMI

Günümüzde sosyal ağlar, insanların birbirleriyle olan etkileşimlerini ve ilişkilerini analiz etmek için önemli bir veri kaynağıdır. Ancak, büyük ve karmaşık sosyal ağ yapılarında, kullanıcıların ilgi alanlarına veya sosyal çevrelerine göre en uygun arkadaşları önermek büyük bir zorluktur. Bu proje, bu zorluğu aşarak, sosyal ağ kullanıcılarına en uygun arkadaş önerilerini sunacak bir sistem geliştirmeyi amaçlamaktadır.

PROBLEMİN AMAÇ VE HEDEFLERİ

Bu projenin temel amacı, sosyal ağ analizi kullanarak, kullanıcılar arasındaki ilişkileri anlamak ve Graph Neural Network (GNN) modeli kullanarak, ağdaki düğümler üzerinden bir öneri sistemi geliştirmektir. Hedefler şunlardır:

* Sosyal ağ verilerini kullanarak, kullanıcılar arasındaki ilişkileri modellemek.
* GNN modeli ile kullanıcıların potansiyel arkadaşlarını tahmin etmek.
* Ağa yeni katılan kullanıcılar için en uygun 10 kişiyi öneren bir sistem oluşturmak.
* Neo4j veri tabanı entegrasyonu ile dinamik veri yönetimi sağlamak.

PROJE KAPSAMI

Bu proje, sosyal ağ analizine odaklanarak, bir öneri sistemi geliştirmeyi kapsamaktadır. Proje, veri setinin içe aktarılması, graf yapısının oluşturulması, GNN modeli ile öneri sistemi geliştirilmesi ve Neo4j veri tabanı entegrasyonunu içermektedir. Projede ayrıca, dinamik olarak yeni düğüm ve kenar ekleme, silme ve kontrol işlemleri de gerçekleştirilecektir.

ÖZGÜN DEĞER

Bu projenin özgün değeri, sosyal ağ analizi ve Graph Neural Network (GNN) modellerinin birleştirilerek, kullanıcılar arasındaki ilişkileri daha doğru ve etkili bir şekilde modelleyebilmesidir. Projenin, sosyal ağlardaki kullanıcılar için daha doğru ve kişiselleştirilmiş arkadaşlık önerileri sunarak, kullanıcı deneyimini artırması hedeflenmektedir. Ayrıca, Neo4j entegrasyonu ile verinin dinamik olarak yönetilmesi, projenin pratikte uygulanabilirliğini artırmaktadır.

KULLANILAN VERİ SETİ

Proje kapsamında kullanılacak veri seti, Stanford Üniversitesi'nin SNAP platformundan alınmıştır ve Facebook verilerini içermektedir. Veri seti, aşağıdaki dosya tiplerinden oluşmaktadır:

* Circles: Ego kullanıcısının çemberlerini içerir.
* Edges: Kullanıcılar arasındaki bağlantıları içerir.
* Egofeat: Ego kullanıcısının özelliklerini içerir.
* Feat: Ağdaki her bir düğümün özelliklerini içerir.
* Featnames: Özellik isimlerini içerir.

KULLANILAN YÖNTEMLER

Proje kapsamında kullanılacak yöntemler şunlardır:

* Veri İçe Aktarımı: Veri setinin içe aktarılması ve ön işleme tabi tutulması.
* Graf Yapısının Oluşturulması: NetworkX kütüphanesi kullanılarak, veri setinden graf yapısının oluşturulması.
* GNN Modelinin Geliştirilmesi: PyTorch Geometric kullanılarak, GNN modelinin oluşturulması ve eğitilmesi.
* Modelin Eğitimi ve Değerlendirilmesi: Modelin K-Fold Cross Validation yöntemi ile eğitilmesi ve performansının değerlendirilmesi.
* Tahmin ve Öneri Sistemi:Yeni kullanıcılar için en uygun arkadaş önerilerini yapacak bir sistem geliştirilmesi.
* Neo4j Entegrasyonu: Graf yapısının Neo4j veri tabanına aktarılması ve dinamik veri yönetimi sağlanması.

KULLANILAN ARAÇLAR

Proje kapsamında kullanılacak araçlar şunlardır:

* Python: Programlama dili.
* Pandas: Veri işleme kütüphanesi.
* NetworkX: Graf oluşturma ve analiz kütüphanesi.
* PyTorch: Derin öğrenme kütüphanesi.
* PyTorch Geometric: GNN modelleri için PyTorch tabanlı kütüphane.
* Neo4j: Graf veri tabanı.

PROJE AKIŞ ŞEMASI

