|  |  |  |  |
| --- | --- | --- | --- |
| Bugünkü staj çalışmam kapsamında, anomalileri tespit eden yapay zekâ sistemimi daha gelişmiş ve verimli hale getirmek için çeşitli iyileştirmeler gerçekleştirdim. Özellikle autoencoder mimarisi ve buna bağlı prompt mühendisliği alanlarında önemli ilerlemeler kaydettim.  Yeni oluşturduğum autoencoder kodu, hem daha esnek veri yapılarıyla çalışabilecek şekilde güncellendi hem de eğitim sonrası üretilen model dosyalarının daha doğru sonuçlar verebilmesi için optimize edildi. Ayrıca LLM modelinden alınan çıktıları daha anlamlı ve doğru hale getirmek adına prompt yapılarımı yeniden tasarladım.  Bu geliştirmelerin ardından, sistemin yanıt süresini iyileştirmek ve genel işlem verimliliğini artırmak için LLM modelini CUDA destekli GPU üzerinden çalıştırma üzerine yoğunlaştım. Bu süreçte, Linux ortamında CUDA çekirdekleri ile LLM modellerini çalıştırma üzerine çeşitli araştırmalar yaptım ve bazı kod bölümlerini başarılı bir şekilde GPU ile paralel çalışır hale getirdim.  Ancak, veri alma ve okuma işlemleri hâlâ CPU tabanlı çalıştığı için bu kısımlarda beklenen hız artışını sağlayamadım. Bu nedenle, Linux ortamında CUDA ile LLM modellerini daha etkin bir şekilde çalıştırmak için uygun yöntemler ve yapılandırmalar üzerine araştırmalar yapmaya devam ediyorum. Amaç, modelin tüm bileşenlerini GPU üzerinden çalıştırarak yanıt süresini minimuma indirmek ve sistem performansını daha da artırmak. | | | |
| **Sayfa No** | **Çalışmanın** | | **KONTROL** |
|  | Konusu :.........................................  ......................................................... | Yapıldığı Tarih  ...../..../202.. | ......................................  ...................................... |