

Çok Dillilik ve Öğrenme Analizi

22040101123 Yesset Yelebayev

22040101139 Diyorjon Ochilov

Özet - çok dilliliğin öğrencilerin akademik başarıları ve öğrenme hızları üzerindeki etkilerini incelemektedir. Farklı dil bilgisine sahip öğrencilerin yer aldığı bir veri seti kullanılarak, bilinen dil sayısının akademik başarı ve öğrenme hızı ile ilişkisi araştırılmaktadır. Proje kapsamında, çok dilli BERT (Bidirectional Encoder Representations from Transformers) modelinden faydalanılarak, öğrenci metinleri üzerinde derin öğrenme teknikleriyle sınıflandırma yapılmış ve başarı analizleri gerçekleştirilmiştir. İstatistiksel analizler ile dil çeşitliliğinin bilişsel performans üzerindeki etkileri değerlendirilmiş ve görselleştirmelerle veriye dair görsel içgörüler elde edilmiştir. Ayrıca, dil yeterliliğinin öğrenme hızına ve kavrayış düzeyine etkisi detaylı bir şekilde incelenmiş ve çok dilli öğrenciler için eğitimsel yöntemlere yönelik uygulanabilir çıkarımlar sunulmuştur. Bu araştırma, çok dilliliğin bilişsel faydalarını anlamaya ve eğitimdeki etkilerini incelemeye katkı sağlamayı amaçlamaktadır.

1. Giriş

Çok dilliliğin öğrencilerin akademik başarıları ve öğrenme hızları üzerindeki etkisini inceleyen bir projeye dair yapılan analizleri sunmaktadır. Proje, öğrencilerin bildikleri dil sayısının, akademik başarıları ve öğrenme hızları ile nasıl ilişkili olduğunu belirlemeye yönelik olarak çeşitli veri analizi yöntemleri kullanmaktadır. Bu rapor, proje kapsamında yapılan veri oluşturma, analizler ve görselleştirme adımlarını içermektedir.

2. Kullanılan Yöntemler

Projede kullanılan başlıca teknolojiler ve yöntemler şunlardır:

- **Python:** Veri işleme, model oluşturma ve görselleştirme işlemleri için Python programlama dili kullanılmıştır.
- **pandas ve numpy:** Veri işleme ve analiz süreçlerinde kullanılmıştır.
- **matplotlib ve seaborn:** Verinin görselleştirilmesi için kullanılmıştır.
- **transformers:** Çok dilli BERT modelinin kullanımı için gerekli kütüphanedir.

- **sklearn:** Verinin eğitim ve test kümelerine bölünmesi, model değerlendirme metriklerinin hesaplanmasında kullanılmıştır.

3. Veri Seti ve Özellikler

Proje için oluşturulan veri seti, 100 öğrencinin aşağıdaki özelliklerini içermektedir:

- **Öğrenci ID:** Öğrencinin benzersiz kimlik numarası.
- **Bildiği Dil Sayısı:** Öğrencinin bildiği dil sayısı (1 ile 4 arasında).
- **Ana Dil:** Öğrencinin ana dili (Almanca, İngilizce, Türkçe).
- **Akademik Ortalama:** Öğrencinin genel akademik başarı not ortalaması.
- **Öğrenme Hızı Puanı:** Öğrencinin öğrenme hızına dair belirli bir puan.

Veri seti oluşturulurken **numpy** ve **pandas** kullanılarak öğrenci özellikleri rastgele belirlenmiş ve bu veriler daha sonra bir **CSV dosyası** olarak kaydedilmiştir.

4. Veri Analizi ve Görselleştirme

4.1. Temel İstatistiksel Analiz

Veri setindeki öğrencilerin akademik başarıları ve öğrenme hızları, bildikleri dil sayısı ile karşılaştırılmıştır. Bu analizde, dil sayısı arttıkça akademik başarı ve öğrenme hızının da arttığı gözlemlenmiştir. İstatistiksel özet ve korelasyon analizi ile dil sayısı ile akademik başarı arasındaki ilişki incelenmiştir.

4.2. Görselleştirme

Görselleştirme işlemleri ile verinin daha anlaşılır hale getirilmesi sağlanmıştır. Aşağıdaki grafikler üretilmiştir:

1. **Dil Sayısı ve Akademik Başarı İlişkisi:** Öğrencilerin bildikleri dil sayısına göre akademik başarılarının değişimini gösteren **boxplot** ve **histogram** grafiklerini içerir.
2. **Öğrenme Hızı ve Dil Sayısı:** Öğrencilerin bildikleri dil sayısına göre öğrenme hızlarının dağılımını gösteren **scatter plot** grafiklerine yer verilmiştir.
3. **Korelasyon Matrisi:** **heatmap** kullanılarak dil sayısı, akademik başarı ve öğrenme hızı arasındaki korelasyonlar görselleştirilmiştir.

4.3. Korelasyon Analizi

Veri setindeki çeşitli öznitelikler arasındaki ilişkiler **Pearson korelasyonu** kullanılarak hesaplanmış ve korelasyon matrisi oluşturulmuştur. Bu analiz ile dil sayısının akademik başarı ve öğrenme hızı ile pozitif bir ilişki içerisinde olduğu gözlemlenmiştir.

Bu durumda, "Çok Dillilik ve Öğrenme Analizi" projenizin raporuna aşağıdaki gibi model kullanımı ve kaynaklar kısmı ekleyebilirsiniz:

5. Kullanılan Modeller ve Yöntemler

Proje kapsamında, öğrencilerin dil bilgisi ile akademik başarıları ve öğrenme hızları arasındaki ilişkileri incelemek için çeşitli derin öğrenme ve makine öğrenme modelleri kullanılmıştır. Projeye dair kullanılan başlıca modeller şunlardır:

Çok Dilli BERT (Bidirectional Encoder Representations from Transformers): Bu model, doğal dil işleme (NLP) görevlerinde yaygın olarak kullanılan bir dil modeli olup, çok dilli veri setleri üzerinde de etkili çalışabilmektedir. "bert-base-multilingual-cased" modelinin kullanımıyla, öğrenci metinleri üzerinde sınıflandırma ve analizler gerçekleştirilmiştir.

Kullanılan Model:

AutoModelForSequenceClassification
(Transformers Kütüphanesi)

Model Özellikleri:

Çok dilli ve bağlamdan bağımsız metin anlayışı

Transformer Tabanlı Eğitim: Öğrenci metinlerinin özelliklerini öğrenmek için transfer öğrenme teknikleri kullanılmıştır. Burada özellikle BERT'in önceden eğitilmiş versiyonları kullanılarak, dilsel özellikler üzerinde daha hassas analizler yapılmıştır.

Sklearn (Scikit-Learn): Verilerin eğitim ve test kümelerine ayrılması, model doğrulama ve değerlendirme aşamalarında kullanılmaktadır. Aynı zamanda doğruluk skorları, sınıflandırma raporları gibi metrikler için de Scikit-Learn yardımcı olmuştur.

6. Sonuçlar

Bu analizde, öğrencilerin bildikleri dil sayısının, hem akademik başarıları hem de öğrenme hızları üzerinde önemli bir etkiye sahip olduğu bulunmuştur. Çok dilliliğin öğrencilerin akademik başarılarını ve öğrenme hızlarını artırabileceği görülmüştür. Bu çalışma, eğitimde çok dilliliğin teşvik edilmesinin öğrenci başarısını iyileştirebileceği fikrini desteklemektedir.

Kaynaklar

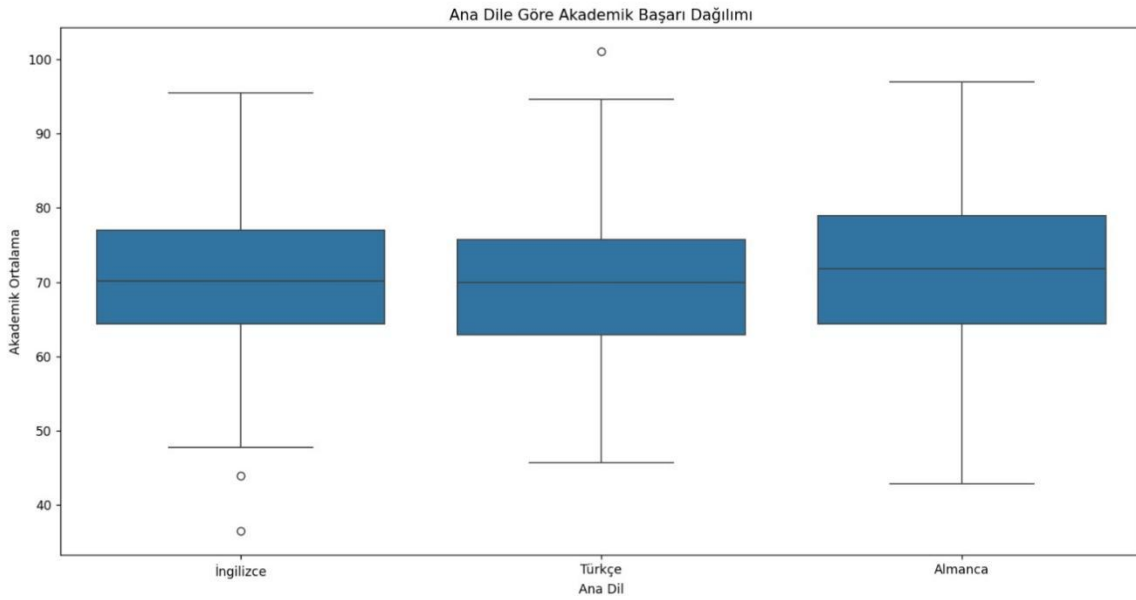
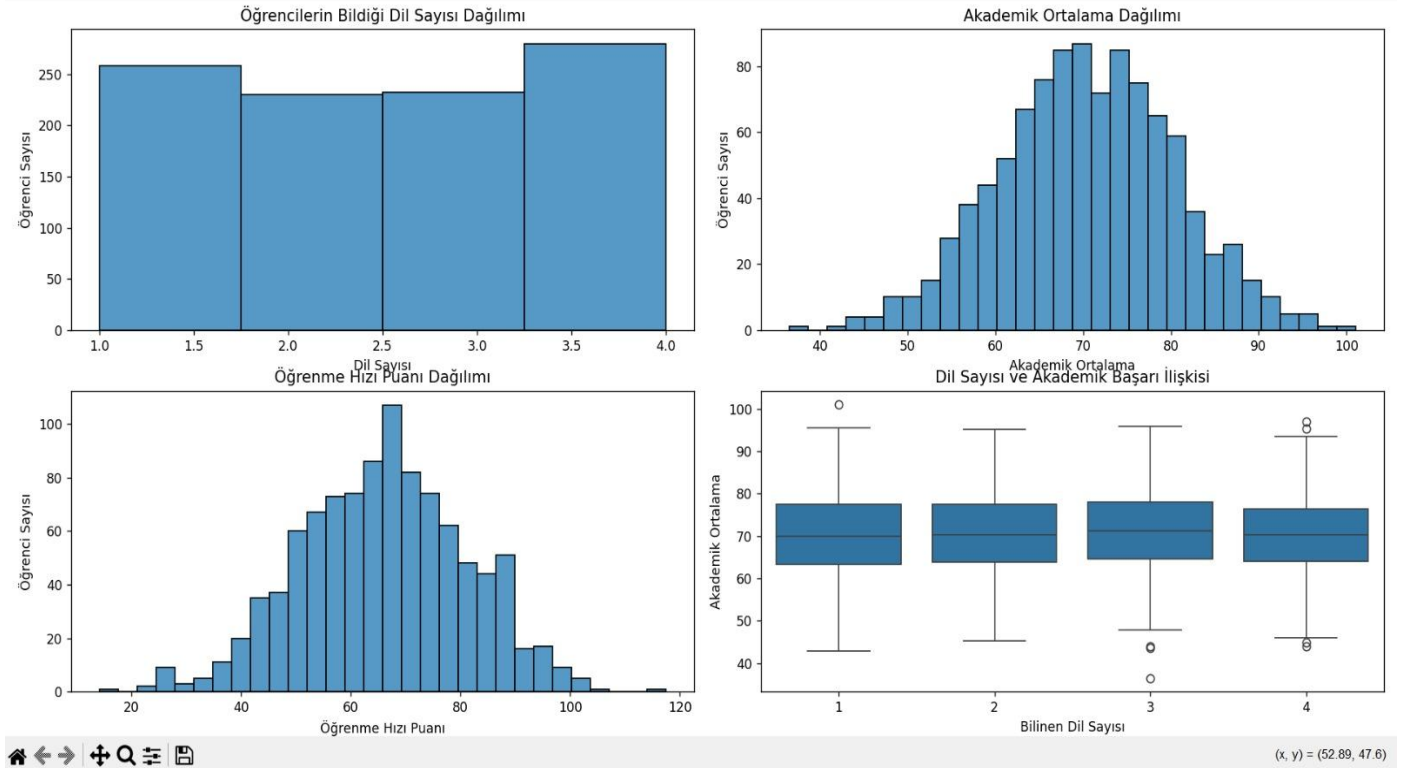
Proje kapsamında kullanılan modeller ve yöntemler ile ilgili daha fazla bilgi edinmek için kaynaklar:

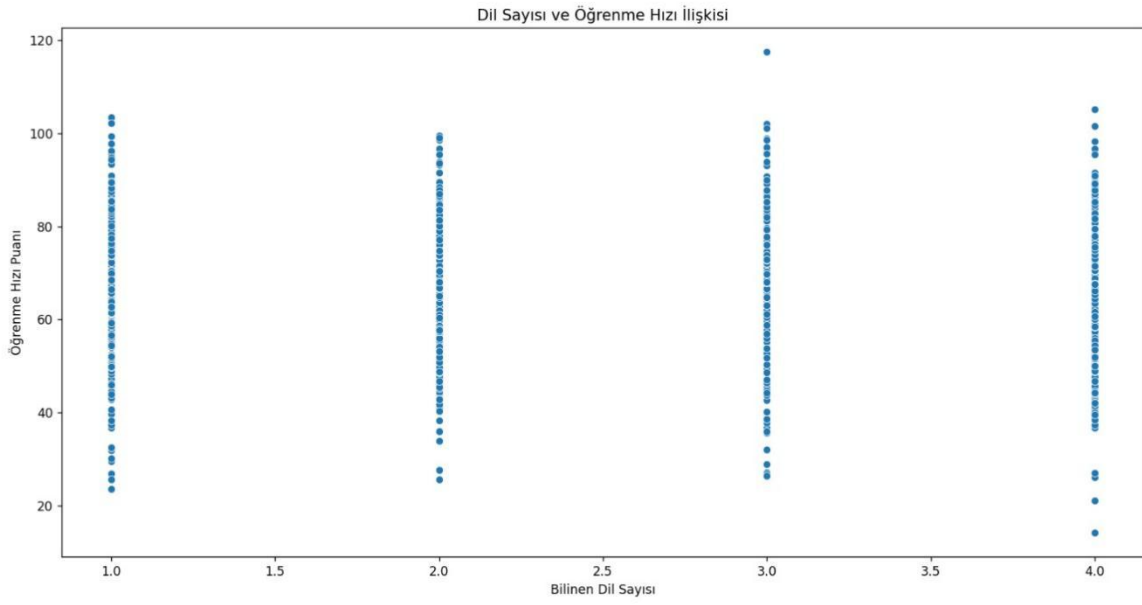
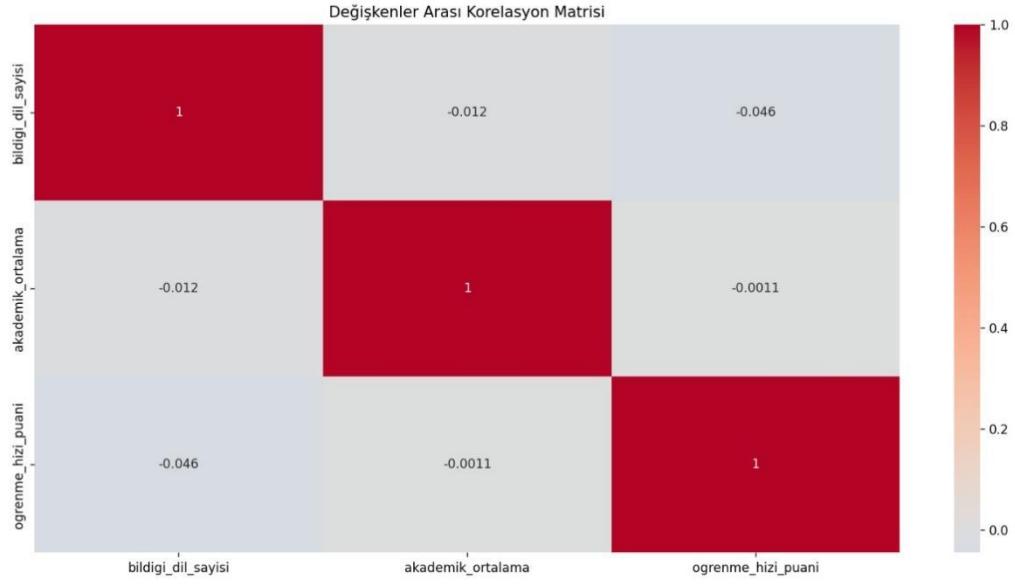
1. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. arXiv:1810.04805.
2. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is All You Need*. arXiv:1706.03762.

3. **Hugging Face (2020).** *Transformers: State-of-the-art Natural Language Processing for Pytorch and TensorFlow 2.0.* Link

4. **Scikit-Learn Documentation. (2020).** *Scikit-Learn: Machine Learning in Python*

[1] .





Diyorjon Ochilov

✉ diyorjonochilov@stu.topkapi.edu.tr ☎ +905425751418 🏠 Istanbul Istanbul Türkiye 🌐 diyor-ochilov



Istanbul Topkapi Üniversitesi'nde eğitim aldım ve henüz iş deneyimim bulunmuyor, ancak stajyerlik için fırsatlar arıyorum. Öğrenme yeteneğim yüksek ve takım çalışmasına yatkınım. Python, HTML ve JavaScript konularında becerilere sahibim. Teknolojiyi yakından takip ederek, kendimi sürekli geliştirmeye özen gösteriyorum. Hedefim, edindiğim teorik bilgileri pratikte uygulayarak profesyonel dünyada sağlam temeller atmak ve deneyim kazanmak. İş yerinize değer katacak bir stajyer olarak katkıda bulunmayı umuyorum.

Eğitimler

2022 - 2026

- **Istanbul Topkapi Üniversitesi**
Bilgisayar Mühendisi, Türkiye

İş deneyimi

- İş deneyimim daha hiç olmadı Stajyerlik için yer arıyorum

Beceriler

Python, C, C++, HTML, CSS, Java,
MsSql, JavaScript

Öğrenme yeteneği, Takım çalışma

Diller

İngilizce

Türkçe

Rusça

Sertifikalar

- **HTML5 ile Web Geliştirme**



YESSET YELEBAYEV

✉ yessetyelebaev@stu.topkapiedu.tr
☎ +905347765029
🏠 İstanbul İstanbul Türkiye

BECERILER

C, C++, Python, Java ,
MySQL, HTML5, CSS3

Ekip çalışması

Öğrenme yeteneği

Araştırma yapma

DİLLER

Türkçe

İngilizce

Rusça

EĞİTİMLER

İstanbul Topkapı Üniversitesi

Eki 2022 - Haz 2026

İŞ DENEYİMİ

Staj için yer arıyorum daha önce hiç iş
deneyimim olmadı