T.C.
SÜLEYMAN DEMİREL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY ZEKANIN LİSE ÖĞRENCİLERİ İÇİN KİMYA VE BİYOLOJİ EĞİTİMİNE ETKİSİ

Pınar KAYA

Danışman
Dr. Öğr. Üyesi Ferdi SARAÇ

YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
ISPARTA - 2025

© 2025 [Pınar KAYA]

İÇİNDEKİLER

Sayfa

İÇİNDEKİLER i
ÖZET iii
ABSTRACT iv
TEŞEKKÜR v
ÇİZELGELER DİZİNİ vi
ŞEKİLLER DİZİNİ vii
SİMGELER VE KISALTMALAR DİZİNİ viii
1. GİRİŞ 1
1.1. Problem Durumu 1
1.2. Araştırmanın Amacı 1
1.3. Araştırmanın Önemi 2
1.4. Araştırma Soruları 2
1.5. Sınırlılıklar 3
2. KURAMSAL ÇERÇEVE VE LİTERATÜR TARAMASI 4
2.1. Yapay Zekâ Kavramı ve Eğitime Yansımaları 4
2.2. ChatGPT ve Gemini'nin Eğitimde Kullanımı 5
2.3. Fen Eğitiminde Yapay Zekâ Kullanımı 5
3. YÖNTEM 8
3.1. Araştırma Modeli 8
3.2. Çalışma Grubu (Evren ve Örneklem) 8
3.3. Veri Toplama Araçları 9
3.4. Verilerin Toplanması ve Analizi 9
4. BULGULAR VE TARTIŞMA 11
4.1. ChatGPT'nin Eğitsel Katkılarına İlişkin Bulgular 12
4.1.1. Farklı Okul Türlerine Göre ChatGPT Ortalama Puanlarının Karşılaştırılması (Kruskal-Wallis Testi Sonuçları) 20
4.1.1.1. Tanımlayıcı İstatistikler (Descriptive Statistics) 21
4.1.1.2. Kruskal-Wallis Mean Rank (Sıra Ortalamaları) 22
4.1.1.3. Chi-Square Test İstatistikleri 22
4.1.1.4. Genel Sonuç ve Yorum 23
4.2. Gemini'nin Eğitsel Katkılarına İlişkin Bulgular 23
4.2.1. Farklı Okul Türlerine Göre Gemini Ortalama Puanlarının Karşılaştırılması (Kruskal-Wallis Testi Sonuçları) 34
4.2.1.1. Tanımlayıcı İstatistikler 34
4.2.1.2. Kruskal-Wallis Mean Rank (Sıra Ortalamaları) 35
4.2.1.3. Kruskal-Wallis Test İstatistikleri 35
4.2.1.4. Genel Sonuç ve Yorum 36
4.3. Ders Bazlı Değerlendirmeler: Biyoloji ve Kimya 37
4.3.1. Farklı Lise Türlerinde Öğrencilerin ChatGPT'nin Biyoloji Dersindeki Kullanımına İlişkin Değerlendirmelerinin Karşılaştırılması 37
4.3.2. Farklı Lise Türlerinde Öğrencilerin ChatGPT'nin Kimya Dersindeki Kullanımına İlişkin Değerlendirmelerinin Karşılaştırılması 38
4.3.3. Farklı Lise Türlerinde Öğrencilerin Gemini'nin Biyoloji Dersindeki Kullanımına İlişkin Değerlendirmelerinin Karşılaştırılması 42
4.3.4. Farklı Lise Türlerinde Öğrencilerin Gemini'nin Kimya Dersindeki Kullanımına İlişkin Değerlendirmelerinin Karşılaştırılması 43
5. SONUÇ VE ÖNERİLER 46

5.1. Sonuçlar 46
5.2. Öneriler 47
KAYNAKLAR 49
ÖZGEÇMİŞ 52

ÖZET

Yüksek Lisans Tezi

YAPAY ZEKANIN LİSE ÖĞRENCİLERİ İÇİN KİMYA VE BİYOLOJİ EĞİTİMİNE ETKİSİ

Pınar KAYA

Süleyman Demirel Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği

Danışman: Dr. Öğr. Üyesi Ferdi SARAÇ

Yapay zekâ modellerinin eğitim üzerindeki artan etkisi akademik araştırmaların kritik bir alanı haline gelmiştir. Büyük dil modelleri, günümüzde yalnızca eğitim sistemleri içinde değil, modern toplumun neredeyse her alanında ölçülebilir etkiler yaratmaktadır.

Bu çalışma üretken yapay zekâ teknolojilerinin lise düzeyinde fen bilimleri eğitimine entegrasyonunu incelemekte olup özellikle kimya ve biyoloji derslerindeki uygulamalarına odaklanmaktadır. Çalışmada ileri düzey dil modellerinin pedagojik kullanım biçimleri araştırılmakta ve bu araçların öğrencilerin kavrama düzeyleri üzerindeki etkileri ile öğretmenlere sunduğu potansiyel destek değerlendirilmektedir. Farklı beceri düzeylerine sahip öğrencilerin ölçülmesini sağlamak amacıyla araştırma farklı lise türlerinde gerçekleştirilmiş ve nitel ile nicel yöntemleri birleştiren karma bir araştırma tasarımı benimsenmiştir. Öğrencilerle yapılan anket sonuçları, bu yapay zekâ araçlarının genel olarak "kısmen yeterli" bulunduğunu göstermektedir. Bununla birlikte bu araçların yeniden kullanılma niyetlerinin okul türüne göre anlamlı biçimde farklılaştığı gözlemlenmiş; bu durum kurumsal bağlamın hem öğrenci kabulünü hem de yapay zekâ destekli eğitim uygulamalarının sürdürülebilirliğini şekillendirmede merkezi bir rol oynadığını ortaya koymaktadır.

Bu çalışmanın büyük dil modellerinin eğitim ortamlarında uygulanmasını inceleyen gelecekteki araştırmalar için bir temel oluşturmasını umuyoruz. Amacımız deneysel içgörüler ve pratik çıkarımlar sunarak bu teknolojilerin öğretim ve öğrenme ortamlarına nasıl etkili bir şekilde entegre edilebileceğine ilişkin sonraki araştırmalara bilgi sağlamak ve rehberlik etmektir.

Anahtar Kelimeler: Yapay Zeka, Eğitim, Lise, Fen Eğitimi, Büyük Dil Modelleri.

2025, 52 sayfa

ABSTRACT

M.Sc. Thesis

THE IMPACT OF ARTIFICIAL INTELLIGENCE ON CHEMISTRY AND BIOLOGY EDUCATION FOR HIGH SCHOOL STUDENTS

Pınar KAYA

Süleyman Demirel University
Graduate School of Natural and Applied Sciences
Department of Computer Engineering

Supervisor: Asst. Prof. Dr. Ferdi SARAÇ

The growing influence of artificial intelligence models on education has become a critical area of academic inquiry. Large language models today exert measurable effects not only within educational systems but also across nearly every domain of modern society.

This study examines the integration of generative artificial intelligence technologies into secondary-level science education, with a particular focus on their applications in chemistry and biology courses. It explores the pedagogical uses of advanced language models and evaluates both their impact on students' conceptual understanding and the potential support they offer to teachers. To ensure the inclusion of students with varying skill levels, the research was conducted across different types of high schools, employing a mixed-methods design that combines qualitative and quantitative approaches. Survey results from students indicate that these AI tools were generally perceived as "partially adequate." However, the intention to reuse these tools differed significantly by school type, suggesting that institutional context plays a central role in shaping both student acceptance and the sustainability of AI-supported educational practices.

We hope that this study will serve as a foundation for future research on the implementation of large language models in educational settings. Our aim is to provide empirical insights and practical implications to inform and guide subsequent investigations into how such technologies can be effectively integrated into teaching and learning environments.

Keywords: Artificial Intelligence, Education, High School, Science Education, Large Language Models.

2025, 52 pages

TEŞEKKÜR

Bu araştırma için beni yönlendiren, karşılaştığım zorlukları bilgi ve tecrübesi ile aşmamda yardımcı olan değerli Danışman Hocam Dr. Öğr. Üyesi Ferdi SARAÇ'a teşekkürlerimi sunarım.

Tezimin her aşamasında beni yalnız bırakmayan aileme sonsuz sevgi ve saygılarımı sunarım.

Pınar KAYA
ISPARTA, 2025

ÇİZELGELER DİZİNİ

Sayfa

Çizelge 4.1. ChatGPT ve Gemini Deneyim Değerlendirme Anketi 11
Çizelge 4.2. ChatGPT Cevaplarının Doğruluk Düzeyi 12
Çizelge 4.3. ChatGPT Cevaplarının Açıklayıcılık Düzeyi 13
Çizelge 4.4. ChatGPT Cevaplarının Anlaşılırlığı 14
Çizelge 4.5. ChatGPT'ye Soru Sorma Kolaylığı 15
Çizelge 4.6. ChatGPT ile Etkileşim Deneyimi 16
Çizelge 4.7. ChatGPT Tekrar Kullanım Eğilimi 18
Çizelge 4.8. ChatGPT'nin Okullara Göre Ortalama Sıra Dağılımı 22
Çizelge 4.9. Chi-Square Testi ile Okullar Arası Fark Analizi 22
Çizelge 4.10. Gemini Cevaplarının Doğruluk Düzeyi 24
Çizelge 4.11. Gemini Cevaplarının Açıklayıcılık Düzeyi 25
Çizelge 4.12. Gemini Cevaplarının Anlaşılırlığı 26
Çizelge 4.13. Gemini'ye Soru Sorma Kolaylığı 27
Çizelge 4.14. Gemini ile Etkileşim Deneyimi 29
Çizelge 4.15. Gemini Tekrar Kullanım Eğilimi 30
Çizelge 4.16. Gemini'nin Okullara Göre Ortalama Sıra Dağılımı 35
Çizelge 4.17. Okullar Arası Gemini Ortalama Puan Farklılıkları 35
Çizelge 4.18. ChatGPT ve Gemininin Okullar Arası Performans Karşılaştırması 36
Çizelge 4.19. Kimya Dersinde ChatGPT_Ortalama Değerlendirmeleri 40
Çizelge 4.20. ChatGPT'nin Kimya Dersindeki Değerlendirmeleri: Okullar Arası Farklılık 40
Çizelge 4.21. İstatistiksel Sonuçlar 41

ŞEKİLLER DİZİNİ

Sayfa

Şekil 4.1. Okul Bazında ChatGPT Ortalama Puanları ve Kruskal-Wallis Sonuçları... 21
Şekil 4.2. Okul Bazında Gemini Ortalama Puanları ve Kruskal-Wallis Sonuçları..... 34
Şekil 4.3. ChatGPT'nin Biyoloji Dersinde Kullanımı: Lise Türlerine Göre Karşılaştırma..... 37
Şekil 4.4. ChatGPT'nin Kimya Dersinde Kullanımı: Lise Türlerine Göre Karşılaştırma..... 39
Şekil 4.5. Gemini'nin Biyoloji Dersinde Kullanımı: Lise Türlerine Göre Karşılaştırma..... 42
Şekil 4.6. Gemini'nin Kimya Dersinde Kullanımı: Lise Türlerine Göre Karşılaştırma..... 44
Şekil 4.7. Okul ve derse göre ChatGPT ve Gemini Ki-Kare Değerleri Karşılaştırması..... 45

SİMGELER VE KISALTMALAR DİZİNİ

AI	Artificial Intelligence
Asymp. Sig	Anlamlılık Düzeyi
df	Serbestlik Derecesi
LLM	Büyük Dil Modelleri
MaaS	Model as a Service
NLP	Doğal Dil İşleme
Npar	Parametrik Olmayan Test
YZ	Yapay Zeka

1. GİRİŞ

1.1. Problem Durumu

Son yıllarda yapay zekâ doğal dil işleme, Web 3.0 (Gan vd., 2023a) ve derin öğrenme alanlarındaki hızlı teknolojik gelişmelerin etkisiyle kayda değer bir ilerleme kaydetmiştir (Xu vd., 2024; Zhao vd., 2020). Büyük dil modelleri (LLM'ler), geniş ölçekli metinleri hem anlamlandırabilen hem de üretebilen yapay zekâ sistemleri olarak öne çıkmaktadır. Mantıksal akıl yürütme yetenekleri sayesinde LLM'ler, özellikle eğitim sektörü başta olmak üzere çeşitli alanlarda hizmet olarak model (MaaS) yaklaşımıyla yaygın biçimde kullanılmaya başlanmıştır (Gan, vd., 2023b).

Son yıllarda özellikle OpenAI tarafından geliştirilen ChatGPT ve Google tarafından geliştirilen Gemini, günümüzde yaygın olarak kullanılan en ileri düzey doğal dil işleme (NLP) teknolojileri arasında yer almaktadır (Yıldız Durak vd., 2025). Bu araçlar yalnızca bilgi sunmakla sınırlı kalmayıp açıklayıcı, yönlendirici ve etkileşimli içerikler üretebilme kapasiteleri sayesinde öğrencilerin öğrenme deneyimlerini zenginleştirmektedir. Eğitim açısından bakıldığında bu modeller öğrencilere bireysel öğrenme fırsatları sunmakta ve kısa ders sürelerinden kaynaklanan eksikliklerini gidermelerine imkân sağlamaktadır (Topal vd., 2021). Ancak bu modellerin eğitimde kullanımına yönelik çalışmalar oldukça sınırlıdır (Aktay vd., 2023; Karabıyık, 2024). Bildiğimiz kadarı ile Türkiye'de lise düzeyinde hem kimya hem de biyoloji eğitimine yapay zekanın etkisini araştıran bir çalışma bulunmamaktadır.

1.2. Araştırmanın Amacı

Bu çalışmanın temel amacı yukarıda belirtilen literatürdeki boşluğu doldurmak, ChatGPT ve Gemini'nin lise düzeyindeki kimya ve biyoloji derslerinde nasıl kullanılabileceğini incelemek, söz konusu teknolojilerin öğrencilerin öğrenme süreçlerine sağladığı katkıları ve karşılaşılan sınırlılıkları ortaya koymaktır. Bir diğer ifade ile yapay zekâ modellerinin lise düzeyinde kimya ve biyoloji eğitimine sağlayabileceği katkıları açığa çıkarmak ve bu kapsamda ChatGPT ile Gemini büyük dil modellerini karşılaştırmalı olarak analiz etmektir. Bu çalışmanın lise eğitiminde

büyük dil modellerinin olası katkılarını ele alacak gelecekteki araştırmalara rehberlik etmesi ve katkı sunması umulmaktadır.

1.3. Araştırmanın Önemi

Eğitim ortamlarında yapay zekâ teknolojilerinin kullanımı pedagojik yaklaşımlarda dönüşüm yaratmakta ve bireyselleştirilmiş öğrenme anlayışını desteklemektedir. Bilhassa kavramsal yoğunluğu yüksek olan kimya ve biyoloji gibi derslerde öğrencilerin öğrenme güçlüklerini aşmalarına yardımcı olan bu araçlar öğretim süreçlerinin etkililiğini artırma potansiyeline sahiptir.

Gemini, ayrıntılı analiz ve derinlemesine inceleme gerektiren problem çözme senaryolarında üstün performans sergilerken ChatGPT hız, bağlamsal anlama, metin işleme ve genel verimlilik açısından öne çıkmaktadır (Firdaus vd., 2024; Rane vd., 2024). Ancak bu araçların lise düzeyindeki fen bilimleri derslerinde pedagojik olarak nasıl kullanılabileceğine dair uygulamalı çalışmaların yetersizliği mevcut araştırmanın önemini artırmaktadır. Bu yönüyle bu çalışma hem literatürdeki boşluğu doldurmayı hem de öğretmenler ile eğitim uygulayıcılarına somut örnekler sunmayı amaçlamaktadır.

1.4. Araştırma Soruları

Bu araştırma aşağıdaki sorulara yanıt aramaktadır:

- ChatGPT ve Gemini lise düzeyinde kimya ve biyoloji derslerinde nasıl etkili kullanılabilir?
- Lise düzeyindeki öğrencilerin bu araçların daha iyi olmasını sağlayacak önerileri var mı?
- ChatGPT ve Gemini'nin eğitimsel katkıları ve sınırlılıkları nelerdir?
- Farklı lise türlerinde (Fen Lisesi, Anadolu Lisesi, Sosyal Bilimler Lisesi) bu araçlara yönelik öğrenci algıları ve deneyimleri arasında anlamlı farklılıklar var mıdır?

Elde edilen veriler öğrencilerin gerçek deneyimlerini yansıtmaktadır.

1.5. Sınırlılıklar

Araştırmanın sınırlılıkları aşağıda belirtilmiştir:

- Çalışma yalnızca Isparta ilindeki üç farklı lise türünde gerçekleştirilmiştir.
- Araştırma kapsamı yalnızca kimya ve biyoloji dersleri ile sınırlandırılmıştır.
- Kullanılan yapay zekâ araçları ChatGPT ve Gemini ile sınırlıdır, diğer yapay zekâ araçları araştırma kapsamına alınmamıştır.
- Elde edilen veriler uygulamaya katılan belirli bir öğrenci grubunun görüşleriyle sınırlıdır.

2. KURAMSAL ÇERÇEVE VE LİTERATÜR TARAMASI

2.1. Yapay Zekâ Kavramı ve Eğitime Yansımaları

Yapay zekâ (YZ) insan benzeri düşünme, öğrenme, problem çözme ve karar verme yeteneklerini taklit edebilen bilgisayar sistemleri olarak tanımlanmaktadır (Russell ve Norvig, 2021; Du-Harpur vd., 2020). Eğitim alanında yapay zekâ teknolojilerinin kullanımı öğrenme ortamlarının dijitalleşmesiyle birlikte önemli bir ivme kazanmıştır. YZ'nın eğitim alanında kullanılması öğrenme ve öğretme süreçlerini optimize edilebilmesi için fırsatlar sunmaktadır (Farazouli vd., 2024; Leite, 2023)

ChatGPT ve Gemini gibi büyük dil modelleri karmaşık problem çözme, mantıksal çıkarım yapma ve bağlamsal ilişkileri analiz etme yetkinliklerinin yanı sıra güçlü soru cevaplama becerileri sergilemektedir (Farquhar vd., 2024). Bu dönüşüm sadece teknik bir yenilik değil aynı zamanda pedagojik tasarım süreçlerini de yeniden yapılandıran bir yaklaşım değişimini temsil etmektedir.

Üretken yapay zekâ mevcut verilerden yola çıkarak yeni ve özgün içerikler (metin, görsel, ses vb.) oluşturabilen algoritmaları kapsamaktadır. Özellikle büyük dil modellerine dayalı üretken YZ sistemleri, doğal dil işleme becerileriyle insan benzeri metinler üretebilmekte (Yıldız Durak, 2023; Brown vd., 2020), açıklama yapma, örnek verme, kavramları yeniden ifade etme gibi üst düzey bilişsel işlevleri destekleyebilmektedir.

Bu sistemler öğrencilerin sorularına yanıt vermekle kalmayıp açıklayıcı içerikler sunmakta, yazılı üretim becerilerini geliştirmekte ve çok çeşitli öğrenme stillerine hitap etmektedir. Eğitimde bu teknolojilerin kullanımı öğrenci motivasyonunu ve katılımını artırmakta öğrenme sürecini daha etkileşimli ve kişiselleştirilmiş hale getirmektedir (Zawacki-Richter vd., 2019). Bununla birlikte üretken YZ'nin eğitimde uygulanması bazı zorlukları da beraberinde getirmektedir. Monib (2024), ChatGPT gibi üretken yapay zekâ sistemlerinin eğitimdeki etik, pedagojik ve teknik sorunlarını ele almakta ve etkili kullanım stratejilerini tartışmaktadır.

2.2. ChatGPT ve Gemini'nin Eğitimde Kullanımı

AI temelli öğrenme sistemlerinin öğrenme deneyimlerini bireysel profillere, çeşitli öğrenme stillerine ve hızlarına uyarlayarak eğitim kalitesini önemli ölçüde artırdığı (Agbong-Coates, 2024; Katiyar, 2024), bunların sonucunda ise öğrenci katılımı ve memnuniyetinin de arttığı belirlenmiştir (Kamruzzaman, 2023; Nurjanah, 2024).

ChatGPT eğitim ortamlarında metin oluşturma, kavram açıklama, yazılı anlatımı geliştirme ve anlık destek sağlama gibi amaçlarla yaygın olarak kullanılmaktadır (Brown vd., 2020). ChatGPT'nin öğrencilerin kavramsal anlama düzeyini artırdığı (Wang ve Fan, 2025), öğrenme sürecine aktif katılım sağladığı ve akademik başarıya olumlu katkılar sunduğu yönünde çeşitli bulgular elde edilmiştir (Fryer ve Dinsmore, 2023). Aynı zamanda ChatGPT, çeşitli doğal dil işleme görevlerinde, örneğin soru cevaplama, makaleler yazma ve çeviri gibi görevlerde son derece etkili performans göstermektedir (Kasneci vd., 2023).

Gemini ise metin, görsel ve sesli içerikleri aynı anda işleyebilme kapasitesi sayesinde farklı öğrenme stillerine hitap edebilmekte ve özellikle görselleştirme ihtiyacının yüksek olduğu alanlarda etkili bir öğrenme aracı olarak kullanılabilmektedir (Imran ve Almusharraf, 2024). Henüz ChatGPT kadar yaygın olmasa da Gemini'nin potansiyelinin yüksek olduğu birçok çalışma tarafından vurgulanmaktadır (Zhou vd., 2024).

Sonuç olarak büyük dil modelleri (LLM'ler) hâlen öğrencilerin ve öğretmenlerin öğrenme ile öğretme süreçlerine katkı sağlama potansiyeline sahiptir, özellikle anında geri bildirim sunma ve kişiselleştirilmiş öğrenme deneyimleri oluşturma konusunda etkili olabilmektedir (Dao vd., 2023). Dolayısıyla eğitim alanında büyük dil modelleri hem öğrenciler hem de öğretmenler için faydalı bir araç olarak kullanılabilir.

2.3. Fen Eğitiminde Yapay Zekâ Kullanımı

Fen bilimleri eğitimi kavramsal yoğunluğu ve soyut kavramlara dayalı içeriği nedeniyle öğrencilerin çok fazla zorlandığı alanlardan biridir. Bu noktada üretken yapay zekâ araçlarının öğrencilerin kavramsal anlayışını desteklemede, soyut

kavramları somutlaştırmada ve bireysel öğrenme süreçlerini güçlendirmede önemli bir rol oynadığı görülmektedir (Huang vd., 2022; Kim, 2023).

Nguyen vd. (2023), Vietnam'daki lise kimya müfredatına ChatGPT'nin entegrasyonunun etkinliğini araştırmıştır. Çalışmada ChatGPT'nin 2019 ve 2023 yılları arasında Vietnam Ulusal Lise Mezuniyet Sınavı'ndaki kimya sorularını çözmedeki performansı değerlendirilmiştir. Bulgular, ChatGPT'nin doğruluğunun yetersiz kaldığını ve öğrencilerin modeli sürekli olarak geride bıraktığını göstermektedir.

Başka bir çalışmada (Kharchenko ve Babenko, 2024) kimya eğitiminde özellikle ChatGPT, Gemini ve Copilot gibi büyük dil modellerinin kullanımı kimyayla ilgili görevleri yerine getirmedeki etkinliklerine odaklanılarak incelenmiştir. Bulgular, bu modellerin derin muhakeme gerektirmeyen problemleri çözmede yeterli olsa da kimya görevlerini tamamlamadaki genel performanslarının öğrencilerinkinden daha düşük olduğunu göstermiştir.

Yakın zamanda yapılan bir çalışmada (de Zárate vd., 2024) araştırmacılar, öğretmenin olmadığı durumlarda kimya ve fizik öğrenimini desteklemek için ChatGPT'nin bir eğitim aracı olarak kullanılmasının etkisini araştırdılar. Valensiya'da 15-16 yaş arası bir grup öğrenciyle yürütülen çalışma, ChatGPT'nin kimya öğrenme çıktılarını nasıl etkilediğini inceledi (de Zárate vd., 2024).

Bir başka çalışmada (Karabıyık, 2024) ChatGPT'nin matematik eğitiminde kullanımını ele alınarak öğrenci ve öğretmen deneyimlerinin değerlendirilmesi amaçlanmıştır. Çalışma dört öğretmen adayı ve on yedi mühendislik öğrencisinden oluşan bir katılımcı grubunu kapsamaktadır. Katılımcılar, ChatGPT ile etkileşimde bulunmuş ve bu etkileşimlere ilişkin veriler analiz edilmiştir. Çalışma sonucunda elde edilen bulgular ChatGPT'nin matematik problemlerinin çözümünde etkili olduğunu belirtmiş buna karşın geometri konularında ve yanlış anlamaların düzeltilmesinde sınırlı kaldığı tespit edilmiştir.

Amasya'da gerçekleştirilen bir başka yakın zamanlı çalışmada (Torun ve Karamustafaoğlu, 2025), 6. sınıf öğrencileri kuvvet konusundaki fen derslerinin bir

parçası olarak toplam on iki saat boyunca ChatGPT ile etkileşimde bulundular. Araştırma, ChatGPT'nin öğrencilerin bu konuyu anlamalarına katkısını incelemeyi amaçladı. Çalışmanın sonunda öğrenciler ChatGPT hakkında olumlu görüşler dile getirdiler ve öğrenme süreçlerinde faydalı ve yararlı bir araç olduğunu düşündüler.

Ancak literatürde bu araçların lise düzeyinde ve özellikle Türkiye bağlamında fen eğitimi alanında nasıl kullanıldığına dair uygulamalı ve karşılaştırmalı araştırmalar oldukça sınırlıdır. Ayrıca bildiğimiz kadarıyla Türkiye'de lise düzeyinde hem kimya hem de biyoloji eğitimine yapay zekânın etkisini inceleyen bir çalışma bulunmamaktadır. Bu çalışmada LLM'lerin lise düzeyindeki eğitime katkıları incelenmiş, öğrenci görüşleri değerlendirilmiş ve modellerin verdiği cevapların doğrulukları belirlenerek birbirleriyle karşılaştırmaları yapılmıştır. Bunlar da çalışmamızın özgünlüğünü ve literatüre katkısını ortaya koymaktadır.

3. YÖNTEM

3.1. Araştırma Modeli

Bu araştırma karma yöntem yaklaşımıyla yürütülmüştür. Nicel ve nitel veri toplama tekniklerinin birlikte kullanılmasıyla, araştırma konusu hem istatistiksel açıdan hem de bağlamsal olarak (sosyal, kültürel, eğitsel veya çevresel koşullar ve bunların etkileri) çok boyutlu şekilde değerlendirilmiştir. Araştırmada aynı zamanda çoklu durum çalışması yöntemi (multiple case study) kullanılmıştır.

Çoklu durum çalışması, nitel araştırma desenlerinden biri olup birden fazla durumu kendi bağlamı içerisinde derinlemesine incelemeyi amaçlamaktadır. Bu çalışmada üç farklı lise türü, Fen Lisesi, Anadolu Lisesi ve Sosyal Bilimler Lisesi, ayrı birer "durum" olarak ele alınmıştır. Bu yaklaşım, üretken yapay zekâ araçlarının farklı okul türlerinde nasıl etki gösterdiğini karşılaştırmalı olarak inceleme imkânı sunmuştur.

Bu yöntemsel tercihler sayesinde hem belirli bağlamlara özgü bulgulara ulaşılmış hem de daha geniş eğitimsel çıkarımlar yapılabilmiştir.

3.2. Çalışma Grubu (Evren ve Örneklem)

Araştırmanın çalışma grubunu Türkiye'nin Isparta ilinde bulunan üç farklı devlet lisesinde öğrenim gören 10. sınıf öğrencileri oluşturmaktadır. Bu liseler:

- Isparta Fen Lisesi
- Şehit Ali İhsan Kalmaz Anadolu Lisesi (ŞAİK)
- Gazi Sosyal Bilimler Lisesi'dir.

Okulların seçiminde öğrencilerin sosyoekonomik ve kültürel çeşitliliği gözetilmiştir. Dolayısıyla yapay zekâ destekli eğitimin farklı öğrenci profilleri üzerindeki etkileri daha dengeli ve karşılaştırmalı biçimde analiz edilebilmiştir.

Araştırma öncesinde Isparta İl Millî Eğitim Müdürlüğünden gerekli resmi izinler alınmış, Süleyman Demirel Üniversitesi Etik Kurulu tarafından etik onay süreci

tamamlanmıştır. Tüm katılımcılar (öğrenci ve velileri) araştırma hakkında bilgilendirilmiş ve yazılı onam formları ile gönüllülük esasına göre çalışmaya dahil edilmiştir.

3.3. Veri Toplama Araçları

Araştırmada hem nicel hem de nitel veri toplama araçları kullanılmıştır:

- Kapalı uçlu anket formları: ChatGPT ve Gemini kullanımına yönelik öğrenci görüşlerini ölçmek amacıyla Likert tipi ölçekler içeren anketler uygulanmıştır. Burada, Kararsızım 0, Yetersiz 1, Kısmen Yeterli 2 ve Yeterli 3 olacak şekilde 0-3 aralığında değişen bir ölçek ifade edilmektedir. Bu ölçekler; doğruluk, açıklık, anlaşılırlık, yeterlilik ve kullanılabilirlik gibi kriterleri içermektedir.
- Açık uçlu sorular: "ChatGPT'nin/Gemini'nin daha iyi olmasını sağlayacak öneriniz var mı? (Opsiyonel)" gibi sorularla öğrencilerin kişisel görüşleri ve önerileri toplanmıştır.
- Gözlem formları: Uygulama sürecinde araştırmacılar ve öğretmenler tarafından yapılan sınıf içi gözlemler kayıt altına alınmıştır.

Bu araçlar sayesinde öğrencilerin hem deneyimsel tutumları hem de düşünsel geri bildirimleri detaylı biçimde elde edilmiştir.

3.4. Verilerin Toplanması ve Analizi

Uygulama süreci öğrencilerin kimya ve biyoloji derslerinde en çok zorlandıkları konuların belirlenmesiyle başlamıştır. Bu konular ilgili branş öğretmenleriyle birlikte seçilmiş, akabinde bu konulara ilişkin çeşitli sorular ChatGPT ve Gemini gibi üretken yapay zekâ araçlarına yöneltilmiştir. Elde edilen yanıtlar sınıf ortamında öğrencilerle birlikte incelenmiş ve tartışılmıştır.

Öğrenciler, yapay zekâ araçlarından alınan yanıtları değerlendirirken doğruluk, açıklık, anlaşılırlık, yeterlilik ve kullanılabilirlik gibi ölçütler üzerinden puanlama yapmış, ayrıca gözlemsel tepkileri de not edilmiştir.

Toplanan nicel veriler, SPSS istatistik yazılımı kullanılarak analiz edilmiştir. Betimsel istatistiklerin yanı sıra okul türleri arasında anlamlı fark olup olmadığını belirlemek amacıyla Kruskal-Wallis H testi gibi çıkarımsal analizler uygulanmıştır.

Nitel veriler içerik analizi yöntemiyle çözümlenmiş; öğrencilerin açık uçlu yanıtları içeriklerine göre sınıflandırılarak yorumlanmıştır. Gözlem notları ve öneriler de benzer şekilde kategorize edilmiştir.

Bu bütüncül analiz yaklaşımı sayesinde hem istatistiksel sonuçlara hem de bağlamsal yorumlara ulaşılmıştır. Elde edilen bulgular farklı okul türleri arasında yapay zekâ destekli öğrenmenin etkililiği, öğrencilerin tutumu ve pedagojik katkılar açısından farklılık gösterdiğini ortaya koymuştur.

4. BULGULAR VE TARTIŞMA

Bu bölümde ChatGPT ve Gemini adlı üretken yapay zekâ araçlarının lise düzeyindeki kimya ve biyoloji derslerinde kullanımına ilişkin elde edilen bulgulara yer verilmiştir. Veriler hem nicel (anket sonuçları, istatistiksel analizler) hem de nitel (öğrenci görüşleri, gözlemler) yöntemlerle toplanmış ve analiz edilmiştir. Bulgular, okul türleri (Fen Lisesi, Anadolu Lisesi, Sosyal Bilimler Lisesi) ve ders alanları (kimya ve biyoloji) açısından ayrı ayrı değerlendirilmiştir. Nicel veriler SPSS ile analiz edilmiş, nitel veriler ise içerik analiziyle temalandırılmıştır. Öğrencilere uyguladığımız anket Çizelge 4.1.'de gösterilmiştir.

Çizelge 4.1. ChatGPT ve Gemini Deneyim Değerlendirme Anketi

Bu anket, ChatGPT ile etkileşim deneyiminizi değerlendirmek ve geliştirmek için hazırlanmıştır. Lütfen her soruya sizin için en uygun seçeneği işaretleyin. 1. ChatGPT'nin verdiği cevapların doğruluğu nasıldı? ● Yeterli ● Kısmen Yeterli ● Yetersiz ● Kararsızım 2. ChatGPT'nin verdiği cevapların açıklayıcılığı nasıldı? ● Yeterli ● Kısmen Yeterli ● Yetersiz ● Kararsızım 3. ChatGPT'nin verdiği cevapların anlaşılırlığı nasıldı? ● Yeterli ● Kısmen Yeterli ● Yetersiz ● Kararsızım	4. ChatGPT'ye soru sormak kolay ve anlaşılır mıydı? ● Yeterli ● Kısmen Yeterli ● Yetersiz ● Kararsızım 5. ChatGPT ile etkileşim eğlenceli ve faydalı mıydı? ● Yeterli ● Kısmen Yeterli ● Yetersiz ● Kararsızım 6. ChatGPT'yi tekrar kullanmayı düşünür müsünüz? ● Evet ● Hayır ● Kararsızım 7. ChatGPT'nin daha iyi olmasını sağlayacak bir öneriniz var mı? (Opsiyonel)

Bu anket, Gemini ile etkileşim deneyiminizi değerlendirmek ve geliştirmek için hazırlanmıştır. Lütfen her soruya sizin için en uygun seçeneği işaretleyin. 8. Gemini'nin verdiği cevapların doğruluğu nasıldı? ● Yeterli ● Kısmen Yeterli ● Yetersiz ● Kararsızım 9. Gemini'nin verdiği cevapların açıklayıcılığı nasıldı? ● Yeterli ● Kısmen Yeterli ● Yetersiz ● Kararsızım 10. Gemini'nin verdiği cevapların anlaşılırlığı nasıldı? ● Yeterli ● Kısmen Yeterli ● Yetersiz ● Kararsızım	11. Gemini'ye soru sormak kolay ve anlaşılır mıydı? ● Yeterli ● Kısmen Yeterli ● Yetersiz ● Kararsızım 12. Gemini ile etkileşim eğlenceli ve faydalı mıydı? ● Yeterli ● Kısmen Yeterli ● Yetersiz ● Kararsızım 13. Gemini'yi tekrar kullanmayı düşünür müsünüz? ● Evet ● Hayır ● Kararsızım 14. Gemini'nin daha iyi olmasını sağlayacak bir öneriniz var mı? (Opsiyonel)

4.1. ChatGPT'nin Eğitsel Katkılarına İlişkin Bulgular

Çizelge 4.2.'te öğrencilerin ChatGPT'nin verdiği cevapların doğruluğuna ilişkin değerlendirmelerinin sonuçları yer almaktadır.

Çizelge 4.2. ChatGPT Cevaplarının Doğruluk Düzeyi

ChatGPT'nin verdiği cevapların doğruluğu nasıldı?			
Yeterli	Kısmen Yeterli	Kararsızım	Yetersiz
48	81	4	9

Çizelge 4.2.'e göre

Genel Olumlu Değerlendirme:

"Yeterli" ve "Kısmen Yeterli" diyenlerin toplamı: 48 + 81 = 129 kişi. Bu durum, katılımcıların yaklaşık %91'inin ChatGPT'nin verdiği cevapların doğruluğunu en azından yeterli veya kısmen yeterli bulduğunu gösteriyor.

Kısmen Yeterli Oranı Yüksek:

81 kişi (yaklaşık %57) "kısmen yeterli" demiş. Bu, öğrencilerin çoğunun bazı doğruluk sorunları fark ettiğini ama genel olarak yanıtları makul seviyede bulduğunu gösteriyor. Yani içerikler çoğunlukla doğru bulunsa bile mutlak doğruluk yerine temkinli bir kabul söz konusu görünüyor.

Olumsuz ve Kararsız Yanıtlar:

"Yetersiz" diyen 9 kişi (yaklaşık %6) ve "Kararsızım" diyen 4 kişi (yaklaşık %3) var. Bu oranlar düşük olsa da ChatGPT'nin tüm öğrenciler için her zaman güvenilir bulunmadığını göstermektedir.

Verilere göre öğrencilerin büyük çoğunluğu ChatGPT'nin verdiği cevapları doğruluk açısından yeterli veya kısmen yeterli bulmuştur. Fakat özellikle "kısmen yeterli" yanıtların fazlalığı öğrencilerin bazı yanıtları eksik, yüzeysel ya da belirsiz bulmuş olabileceğini düşündürmektedir. Bu durum ChatGPT'nin doğruluk oranının genel olarak yüksek ama geliştirilmeye açık olduğunu göstermektedir.

Çizelge 4.3.'te öğrencilerin ChatGPT'nin verdiği cevapların açıklayıcılık düzeylerini değerlendirmelerinin sonuçları gösterilmiştir.

Çizelge 4.3. ChatGPT Cevaplarının Açıklayıcılık Düzeyi

ChatGPT'nin verdiği cevapların açıklayıcılığı nasıldı?			
Yeterli	Kısmen Yeterli	Kararsızım	Yetersiz
72	63	4	3

Çizelge 4.3.'e göre

Genel Memnuniyet Oranı:

"Yeterli" ve "Kısmen Yeterli" diyenlerin toplamı: 72 + 63 = 135 kişi. Bu, tüm katılımcıların yaklaşık %95'inin ChatGPT'nin açıklayıcılığını olumlu bulduğunu gösteriyor.

Kısmen Yeterli Cevapların Ağırlığı**:**

63 öğrenci (yaklaşık %44) "Kısmen yeterli" yanıtını vermiştir. Bu katılımcılar ChatGPT'nin genel olarak açıklayıcı olduğunu düşünmekle birlikte bazı içerikleri teknik, uzun ya da karmaşık bulmuş olabilir. Bu sonuç içeriklerin bazı öğrenciler için daha sade, görselle desteklenmiş ya da örnekle zenginleştirilmiş biçimde sunulması gerektiğini ortaya koymaktadır.

Olumsuz Geri Bildirim:

"Yetersiz" diyen sadece 3 kişi var (yaklaşık %2).

"Kararsızım" diyenlerin sayısı da çok düşük: 4 kişi (%3).

Katılımcıların büyük çoğunluğu ChatGPT'nin verdiği cevapları açıklayıcı ve yeterli bulmuştur.

"Kısmen yeterli" oranı da oldukça yüksek olduğundan bazı kullanıcılar daha fazla detay veya açıklama bekliyor olabilir. Ancak genel olarak ChatGPT'nin eğitsel anlamda açıklayıcılığı yüksek bulunmuştur.

Çizelge 4.4.'te ChatGPT'nin sorulara verdiği cevapların anlaşılırlığının öğrenciler tarafından değerlendirilmesinin sonuçları gösterilmiştir.

Çizelge 4.4. ChatGPT Cevaplarının Anlaşılırlığı

ChatGPT'nin verdiği cevapların anlaşılırlığı nasıldı?			
Yeterli	Kısmen Yeterli	Kararsızım	Yetersiz
45	79	6	12

Çizelge 4.4.'e göre

Olumlu Değerlendirmeler:
"Yeterli" ve "Kısmen Yeterli" yanıtları toplamda 45 + 79 = 124 kişi. Bu da katılımcıların yaklaşık %87'sinin ChatGPT'nin açıklamalarını tamamen ya da kısmen anlaşılır bulduğunu göstermektedir.

Kısmen Yeterli Cevapların Ağırlığı:
"Kısmen yeterli" diyen 79 kişi (yaklaşık %56), ChatGPT'nin ifadelerini genel olarak anladıklarını ancak yer yer karmaşık, belirsiz veya teknik bulmuş olabileceklerini düşündürüyor. Bu durum içeriklerin bazı öğrenciler için sadeleştirilmesi gerektiğini ortaya koyabilir.

Olumsuz ve Kararsız Yanıtlar:
"Yetersiz" diyen 12 kişi (yaklaşık %9) ve "Kararsızım" diyen 6 kişi (yaklaşık %4), azınlıkta kalmakla birlikte dikkat çekicidir. Bu kullanıcılar için içerikler muhtemelen daha açık, öğrenci düzeyine uygun ve sade şekilde sunulursa daha etkili olacaktır.

Öğrencilerin büyük çoğunluğu ChatGPT'nin yanıtlarını anlaşılır ya da kısmen anlaşılır bulmaktadır. Ancak "kısmen yeterli" yanıtlarının yüksekliği, dil düzeyi, terim kullanımı veya açıklama tarzının her öğrenci profiline uygun olmadığını göstermektedir. Bu bulgu özellikle ortaöğretim düzeyinde sadeleştirilmiş ve yapılandırılmış anlatımların daha faydalı olabileceğine işaret etmektedir.

Çizelge 4.5.'te öğrencilerin ChatGPT'ye soru sorma kolaylığı hakkındaki görüşlerinin sonuçları gösterilmiştir.

Çizelge 4.5. ChatGPT'ye Soru Sorma Kolaylığı

ChatGPT'ye soru sormak kolay ve anlaşılır mıydı?			
Yeterli	Kısmen Yeterli	Kararsızım	Yetersiz
60	56	11	15

Çizelge 4.5.'e göre

Olumlu Değerlendirme:
"Yeterli" ve "Kısmen Yeterli" diyen toplam 60 + 56 = 116 kişi, bu da toplamın %81,7'sine karşılık gelmektedir. Bu oran öğrencilerin büyük çoğunluğunun ChatGPT'ye soru sorma sürecini kolay ve anlaşılır bulduğunu göstermektedir.

Kısmen Yeterli Oranının Yüksekliği:
"Kısmen yeterli" diyenlerin oranı %39,4 oldukça yüksektir. Bu durum bazı öğrencilerin ChatGPT'ye nasıl soru sorması gerektiğini tam olarak bilememesi ya da daha açıklayıcı bir yönlendirme beklemesiyle ilgili olabilir.

Olumsuz ve Kararsız Yanıtlar:
"Yetersiz" diyen 15 kişi (%10,6), "Kararsızım" diyen 11 kişi (%7,7) ile değerlendirildiğinde yaklaşık %18'lik bir kesim soru sorma sürecinde zorluk yaşamış olabilir. Bu durum özellikle bazı öğrencilerin dijital okuryazarlık düzeyi, deneyim eksikliği veya sistemin beklentilerini anlamada zorlanmalarıyla açıklanabilir.

Öğrencilerin önemli bir kısmı ChatGPT'ye soru sormayı kolay ve anlaşılır bulmaktadır. Ancak "kısmen yeterli" ve "yetersiz" yanıtların oranı, kullanıcı arayüzünün daha yönlendirici olması, örnek sorular verilmesi veya dil kullanımının çeşitlendirilmesi gerektiğini göstermektedir. Özellikle eğitim amaçlı kullanımda öğrencilere ChatGPT ile etkili iletişim kurmaları için rehberlik edilmesi faydalı olabilir.

Çizelge 4.6.'da öğrencilerin ChatGPT ile etkileşimlerinin yeterliliğinin değerlendirilmesi sonuçları belirtilmiştir.

Çizelge 4.6. ChatGPT ile Etkileşim Deneyimi

ChatGPT ile etkileşim eğlenceli ve faydalı mıydı?			
Yeterli	Kısmen Yeterli	Kararsızım	Yetersiz
45	79	6	12

Çizelge 4.6.'ya göre

Olumlu Değerlendirmeler:
"Yeterli" ve "Kısmen Yeterli" diyen toplam 45 + 79 = 124 kişi. Bu, katılımcıların yaklaşık %87,3'ünün ChatGPT ile etkileşimi en azından kısmen faydalı ve eğlenceli bulduğunu göstermektedir.

Kısmen Yeterli Oranının Yüksekliği:
79 öğrenci (%55,6) etkileşimi "kısmen yeterli" olarak değerlendirmiştir. Bu oran etkileşimin bazı yönlerden ilgi çekici ya da verimli olduğunu ancak daha fazla etkileşim veya çeşitlilik beklendiğini gösterebilir.

Olumsuz ve Kararsız Yanıtlar:
"Yetersiz" diyen 12 kişi (%8,5) ve "Kararsızım" diyen 6 kişi (%4,2) ile toplam %12,7'lik bir kesim etkileşimin yeterince eğlenceli veya faydalı olmadığını belirtmiştir. Bu durum, bazı öğrenciler için ChatGPT'nin sunum biçiminin daha etkileşimli, kişiselleştirilmiş ya da yaratıcı hale getirilmesinin gerekli olabileceğine işaret eder.

Katılımcıların büyük çoğunluğu ChatGPT ile kurdukları etkileşimi genel olarak faydalı ve eğlenceli bulmuştur. Ancak "kısmen yeterli" yanıtların yüksekliği bazı kullanıcıların daha etkileyici veya motive edici bir deneyim beklentisi içinde olduklarını göstermektedir. Bu bağlamda ChatGPT'nin öğrenci merkezli, oyunlaştırılmış ya da görsel-işitsel içeriklerle desteklenen sürümlerinin etkileşimi daha da artırabileceği düşünülmektedir.

Araştırmaya katılan 142 öğrenciden elde edilen anket verileri ChatGPT'ye ilişkin genel olarak olumlu bir tutum sergilendiğini göstermektedir. Öğrencilerin yaklaşık %90'ı ChatGPT'yi tekrar kullanmak istediklerini belirtmiştir. Doğruluk, açıklayıcılık, anlaşılırlık, yeterlilik ve kullanılabilirlik kriterlerinde de yüksek düzeyde memnuniyet ifade edilmiştir. Çizelge 4.7'de öğrencilere yöneltilen ChatGPT'yi tekrar kullanmayı düşünür müsünüz sorusundan elde edilen sonuçlar gösterilmiştir.

Çizelge 4.7. ChatGPT Tekrar Kullanım Eğilimi

ChatGPT'yi tekrar kullanmayı düşünür müsünüz?		
Evet	Hayır	Kararsızım
129	1	12

Çizelge 4.7'ye göre

Olumlu Görüşler:
"Evet" diyen katılımcı sayısı 129 kişi, bu da katılımcıların yaklaşık %91'inin ChatGPT'yi yeniden kullanmayı düşündüğünü belirtmektedir.

Olumsuz Görüşler:
"Hayır" diyen yalnızca 1 kişi var. Bu, %1'in altında bir orana denk geliyor ve oldukça düşüktür.

Kararsızlar:
"Kararsızım" diyen 12 kişi, bu da katılımcıların yaklaşık %8.5'ine karşılık gelmektedir. Bu katılımcılar muhtemelen karar vermek için daha fazla deneyime veya bilgiye ihtiyaç duyuyor olabilir.

Veriler ChatGPT'nin 10. sınıf öğrencileri arasında oldukça yüksek bir yeniden kullanım potansiyeline sahip olduğunu ortaya koymaktadır.

Katılımcıların %90'ından fazlasının tekrar kullanmayı düşünmesi, ChatGPT'nin faydalı, pratik ve kullanıcı dostu bir araç olarak değerlendirildiğine işaret etmektedir. "Kararsız" grubun varlığı geliştirme sürecinde farklı kullanıcı beklentilerinin dikkate alınması gerektiğini de gösterir.

Katılımcılara "ChatGPT'nin Daha İyi olmasını Sağlayacak Bir Öneriniz Var Mı?" sorusu opsiyonel olacak şekilde yöneltilmiştir. Sonuçlar lise öğrencilerinin üretken yapay zekâ araçlarından ChatGPT'ye yönelik deneyimlerini ve geri bildirimlerini içermektedir. Öğrencilere ChatGPT'nin kullanım sürecinde karşılaştıkları olumlu ve olumsuz durumları içeren açık uçlu sorular yöneltilmiş, yanıtlar tematik analiz yöntemiyle çözümlenmiştir.

Çizim ve Görsel Destek Başarısı

Öğrenciler ChatGPT'nin çizimlerini genel olarak doğru, başarılı ve görsel açıdan açıklayıcı bulmuştur. Özellikle biyoloji ve kimya derslerindeki organik yapılar ve çizimler beğenilmiştir. Yorumlardan bazıları: "Çizimler çok güzel", "Doğru çiziyor", "Görsel destek iyi" vb.

Açıklayıcı ve Tam Yanıtlar

ChatGPT'nin verdiği yanıtların büyük kısmı öğrencilerce açıklayıcı, anlaşılır ve yeterli düzeyde olarak değerlendirilmiştir. Özellikle konunun özetlenmesi ve tablo ile sunulması memnuniyet yaratmıştır. Öğrencilerin büyük bir kısmı ChatGPT'nin Gemini'ye kıyasla daha başarılı olduğunu ifade etmiştir.

Öğrencilerin büyük çoğunluğu ChatGPT'yi yeniden kullanmak istediklerini ifade etmiş; ayrıca eğlenceli ve etkileşimli yapısının öğrenme üzerinde olumlu etkiler yarattığını belirtmiştir.

Olumsuz Geri Bildirimler ve Gelişim Alanları

Hız ve Performans Sorunları

Görsel oluşturma ve yanıt üretme süreci öğrenciler tarafından yavaş bulunmuştur. Bazı öğrenciler, görsellerin hiç gelmediğini veya eksik çizildiğini belirtmiştir. Öğrenciler, "Çok yavaş çalışıyor", "Fotoğraf geç geliyor", "Bekletiyor" gibi yorumlarda bulunmuşlardır.

Üyelik Sınırlamaları

ChatGPT'nin ücretsiz versiyonunda yer alan üç hak kullanım sınırı öğrenciler tarafından yetersiz bulunmuş, bu durum onların uygulamadan tam anlamıyla yararlanmasını engellemiştir. Ayrıca belirli özelliklere erişim için üyelik zorunluluğu getirilmesi öğrencilerin kullanım sürecini hem teknik açıdan zorlaştırmakta hem de erişilebilirliklerini sınırlamaktadır.

Görsel Açıklama Yetersizlikleri

Bazı çizimler açıklayıcı olmaktan uzak veya eksik detaylıdır. Görsellerin daha hızlı, renkli ve katmanlı olması gerektiği önerilmiştir.

Cevap Hataları ve Aşırı Uzunluk

Yanıtlar zaman zaman hatalı veya konuyla ilgisiz içerikler barındırmaktadır. Gereğinden uzun, detaysız veya müfredat dışı bilgiler dikkat çekmiştir. Öğrenciler, "Çok uzun anlatıyor, sadelik yok", "Fazla akademik anlatıyor" gibi yorumlarda bulunmuşlardır.

Dilde Uygunluk Sorunu

ChatGPT zaman zaman lise seviyesini aşan karmaşık cümleler kullanabilmektedir. Öğrenciler yanıtların daha sade, kısa ve doğrudan olmasını tercih etmektedir.

Kişiselleştirme Eksikliği

Öğrenciler; yaş, sınıf seviyesi ya da mesleki ilgi alanlarına göre cevaplandırma özelliği talep etmiştir. Bu durum yapay zekânın daha esnek ve kişisel hale getirilmesini zorunlu kılmaktadır.

Elde edilen öğrenci geri bildirimleri ChatGPT'nin lise düzeyinde fen eğitimi bağlamında oldukça etkili, açıklayıcı ve kullanıcı dostu bulunduğunu göstermektedir. Fakat özellikle hız, görsel üretim süreci, dil sadeliği ve kişiselleştirme vb. alanlarda iyileştirmeler yapılması gerektiği çok net bir şekilde göstermektedir.

Yapay zekâ destekli araçların öğrenci düzeyine göre yeniden yapılandırılması sade, anlaşılır ve etkileşimli hale getirilmesi gelecekteki eğitim uygulamaları için kritik öneme sahiptir.

4.1.1. Farklı Okul Türlerine Göre ChatGPT Ortalama Puanlarının Karşılaştırılması (Kruskal-Wallis Testi Sonuçları)

Bu alt bölümde farklı okullarda ChatGPT kullanımı sonucunda elde edilen veriler karşılaştırılmıştır. Bu bağlamda Şekil 4.1.'de okul bazında ChatGPT'nin sıra ortalama puanları ile Kruskal-Wallis test sonuçlarına ait bulgular sunulmuştur. Burada gruplar farklı lise türleri ve farklı öğrencilerden olduğu için bağımsızdır, veriler normal dağılıma sahip olmayacağı için Kruskal-Wallis testi için uygundur.

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
ChatGPT_Ortalama	142	2,1749	0,45017	0,67	2,83
Okul Adı	142	6,00	0,842	0,842	7

Kruskal-Wallis Test

Ranks

	Okul Adı	N	Mean Rank
ChatGPT_Ortalama	ŞAİK AL	50	85,10
	Gazi Sosyal Bilimler Lisesi	42	64,05
	Fen Lisesi	50	64,16
	Total	142	

Test Statistics[a,b]

	ChatGPT_Ortalama
Chi-Square	8,559
Df	2
Asymp. Sig.	0,014

a. Kruskal-Wallis Test
b. Grouping Variable: Okul Adı

Şekil 4.1. Okul Bazında ChatGPT Ortalama Puanları ve Kruskal-Wallis Sonuçları

Bu araştırmada okullar arasında ChatGPT'ye yönelik anlamlı farklılıkların varlığını ortaya koymak amacıyla parametrik olmayan test yöntemlerinden biri olan NPar testi uygulanmıştır.

Veriler sıklıkla Likert tipi ("Yeterli", "Kısmen Yeterli", "Yetersiz", "Kararsızım" gibi dereceli) ölçeklerden elde edilmiş olup sayısal değil sıralı (ordinal) düzeydedir. Bu tür verilerde normal dağılım varsayımı genellikle karşılanmadığından normal dağılım varsayımı aramayan nonparametrik testlerin kullanılması tercih edilmiştir.

4.1.1.1. Tanımlayıcı İstatistikler (Descriptive Statistics)

- ChatGPT_Ortalama: Öğrencilerin ChatGPT kullanımına yönelik sıra ortalama puanları 2,17 düzeyindedir. Puanlar 0,67 ile 2,83 arasında değişmektedir.

- Okul Adı**:** Sayısal kodlarla temsil edilen üç okul türü arasında dağılım yapılmıştır. (5 = ŞAİK AL, 6 = Gazi Sosyal Bilimler Lisesi, 7 = Fen Lisesi)

4.1.1.2. Kruskal-Wallis Mean Rank (Sıra Ortalamaları)

Çizelge 4.8.'de ChatGPT'nin okullara göre ortalama sıra dağılımı verilmiştir.

Çizelge 4.8. ChatGPT'nin Okullara Göre Ortalama Sıra Dağılımı

Okul Adı	N	Ortalama Sıra (Mean Rank)
ŞAİK Anadolu Lisesi	50	85,10
Gazi Sosyal Bilimler Lisesi	42	64,05
Fen Lisesi	50	64,16

Çizelge 4.8. dikkate alındığında;

- ŞAİK Anadolu Lisesi öğrencileri ChatGPT'yi diğer okul türlerine kıyasla daha yüksek sıra ortalamasına sahiptir.
- Gazi Sosyal Bilimler Lisesi ve Fen Lisesi öğrencilerinin ortalama sıraları birbirine oldukça yakındır.

4.1.1.3. Chi-Square Test İstatistikleri

Çizelge 4.9.'da Chi-Square testi ile okullar arasındaki farkların analiz sonuçları gösterilmiştir.

Çizelge 4.9. Chi-Square Testi ile Okullar Arası Fark Analizi

Test	Değer
Chi-Square	8,559
Serbestlik Derecesi (df)	2
Anlamlılık Düzeyi (Asymp. Sig.)	0,014

Anlamlılık düzeyi $p = 0{,}014 < 0{,}05$ olduğundan istatistiksel olarak anlamlı bir fark vardır. Yani öğrencilerin ChatGPT'ye ilişkin puanları okul türüne göre değişmektedir. Tüm okullarda aynı şekilde değerlendirilmemiştir.

4.1.1.4. Genel Sonuç ve Yorum

Bu alt bölümde ChatGPT ve Gemini'nin farklı açılarda karşılaştırılmış, karşılaştırma sonuçları detaylı olarak incelenmiştir.

- Kruskal-Wallis testi sonucuna göre okul türleri açısından ChatGPT'ye verilen puanlar arasında anlamlı bir farklılık bulunmaktadır ($p = 0,014$). En yüksek puan ŞAİK Anadolu Lisesi öğrencilerinden gelmiştir.
- Sosyal Bilimler Lisesi ve Fen Lisesi öğrencileri ChatGPT'yi daha benzer ve daha düşük puanlarla değerlendirmiştir.
- ChatGPT'ye yönelik olumlu tutum özellikle Şehit Ali İhsan Kalmaz Anadolu Lisesi (ŞAİK AL) öğrencileri arasında daha belirgindir. SPSS analizlerine göre bu okulun öğrencileri ChatGPT'yi diğer okul türlerine kıyasla daha olumlu değerlendirmiştir ($p = 0.014<0.05$). Bu sonuç okul türüne göre öğrencilerin üretken yapay zekâ araçlarını algılamalarında anlamlı farklılıklar olduğunu göstermektedir.
- Veriler tüm okul türlerinde öğrencilerin büyük çoğunluğunun ChatGPT'yi tekrar kullanmayı düşündüğünü göstermektedir:

Bu sonuç ChatGPT'nin özellikle doğruluk, açıklayıcılık ve anlaşılırlık gibi akademik kriterlerde öğrencilerin beklentilerini yüksek düzeyde karşıladığını göstermektedir. Fen Lisesi öğrencilerinin bilimsel doğruluğa önem vermesine rağmen ChatGPT'yi yüksek oranda benimsemiş olmaları dikkat çekicidir. Sosyal Bilimler Lisesi öğrencileri ise özellikle kavramsal açıklamalar bakımından ChatGPT'yi olumlu değerlendirmiştir.

4.2. Gemini'nin Eğitsel Katkılarına İlişkin Bulgular

Çizelge 4.10.'da Gemini'nin verdiği yanıtların doğruluğunun öğrenciler tarafından yapılan değerlendirme sonuçları sunulmuştur.

Çizelge 4.10. Gemini Cevaplarının Doğruluk Düzeyi

Gemini'nin verdiği cevapların doğruluğu nasıldı?			
Yeterli	Kısmen Yeterli	Kararsızım	Yetersiz
19	71	7	45

Çizelge 4.10.'a göre

Olumlu Değerlendirmeler:

- "Yeterli" ve "Kısmen Yeterli" yanıtlarının toplamı: 19 + 71 = 90 kişi (yaklaşık %63). Bu oran öğrencilerin yaklaşık üçte ikisinin Gemini'nin verdiği cevapları tam ya da kısmen doğru bulduğunu göstermektedir.

Kısmen Yeterli Oranı Çok Yüksek:

- "Kısmen yeterli" diyen 71 kişi (%50) Gemini'nin yanıtlarının çoğu zaman doğru ama kimi zaman da eksik, yüzeysel veya belirsiz bulunduğunu düşündürmektedir. Bu durum öğrencilerin doğruluk konusunda güvensizlik yaşadıklarını ama arada doğru içeriklere de rastladıklarını gösterir.

Olumsuz Görüşler:

- "Yetersiz" diyen 45 kişi (yaklaşık %32) ile dikkat çekici bir orana sahiptir. Bu katılımcılar Gemini'nin cevaplarını büyük oranda hatalı, yanlış yönlendirici veya eğitsel açıdan yetersiz bulmuştur.

Kararsızlar:

- 7 öğrenci (yaklaşık %5) Gemini'nin doğruluğu konusunda kararsızdır, bu da araçla sınırlı etkileşim ya da belirsiz deneyim yaşandığını belirtir.

Veriler Gemini'nin verdiği yanıtların doğruluğu konusunda öğrencilerin önemli bir kısmının temkinli veya olumsuz bir bakışa sahip olduğunu göstermektedir. Yalnızca %13,4'lük bir kesim "doğruluğu yeterli" bulurken %31,7'si açıkça yetersiz bulmaktadır.

En yüksek oran %50 ile "kısmen yeterli" seçeneğinde yoğunlaşmış olup bu durum Gemini'nin güvenilirliği ve içerik kalitesi bakımından geliştirilmeye açık olduğunu

ortaya koymaktadır. Eğitim ortamlarında kullanılacak bir yapay zekâ aracının doğruluk konusunda daha tutarlı ve güvenilir olması gerektiği açıktır.
Çizelge 4.11.'de Gemini'nin verdiği cevapların açıklayıcılık düzeyinin öğrenciler tarafından değerlendirilmesi sonuçları verilmiştir.

Çizelge 4.11. Gemini Cevaplarının Açıklayıcılık Düzeyi

Gemini'nin verdiği cevapların açıklayıcılığı nasıldı?			
Yeterli	Kısmen Yeterli	Kararsızım	Yetersiz
24	80	6	32

Çizelge 4.11.'e göre

Olumlu Değerlendirmeler:

- "Yeterli" ve "Kısmen Yeterli" yanıtları toplamda 24 + 80 = 104 kişi. Bu, katılımcıların yaklaşık %73'ü Gemini'nin verdiği yanıtları tamamen ya da kısmen açıklayıcı bulduğunu göstermektedir.

Kısmen Yeterli Oranının Ağırlığı:

- En yüksek frekans 80 öğrenciyle (yaklaşık %56) "kısmen yeterli" seçeneğindedir. Bu durum öğrencilerin çoğunun Gemini'nin yanıtlarını genel olarak anlaşılır bulduğunu ancak açıklamaların derinliğinin yetersiz veya yüzeysel kaldığını düşündüğünü göstermektedir.

Olumsuz Görüşler

- "Yetersiz" diyen 32 öğrenci yani yaklaşık %23'ü Gemini'nin açıklamalarını anlamayı zorlaştıracak kadar eksik veya kafa karıştırıcı bulmuştur. Bu oran açıklayıcılık konusunda bazı önemli eksikliklere işaret etmektedir.

Kararsızlar:

- 6 kişi (yaklaşık %4) kararsız olduğunu belirtmiş. Bu, genellikle sınırlı deneyim ya da belirgin bir olumlu/olumsuz yönelim görmeyen kullanıcıları yansıtmaktadır.

Katılımcıların yaklaşık dörtte üçü Gemini'nin verdiği cevapları tam ya da kısmen açıklayıcı bulmuştur. Ancak açıklayıcılığın büyük oranda "kısmen yeterli" düzeyde kalması ve "yetersiz" bulanların oranının da %22'yi geçmesi Gemini'nin içerik sunma şeklinin daha sistemli, örnekli, sade ve anlaşılır hale getirilmesi gerektiğini ortaya koymaktadır. Özellikle lise öğrencileriyle yapılan bir çalışmada açıklayıcılık düzeyinin öğrenci seviyesine uygun, açık ve yapılandırılmış olması kritik düzeyde önemdedir.

Çizelge 4.12.'te Gemini' nin verdiği cevapların anlaşılırlığının öğrenciler tarafından test edilmesinin sonuçları verilmiştir.

Çizelge 4.12. Gemini Cevaplarının Anlaşılırlığı

Gemini' nin verdiği cevapların anlaşılırlığı nasıldı?			
Yeterli	Kısmen Yeterli	Kararsızım	Yetersiz
14	75	7	46

Çizelge 4.12.'ye göre

Olumlu Değerlendirmeler:

- "Yeterli" ve "Kısmen Yeterli" yanıtları toplamda 14 + 75 = 89 kişi tarafından verilmiştir. Bu, katılımcıların yaklaşık %63'ünün Gemini'nin verdiği yanıtları tamamen ya da kısmen anlaşılır bulduğunu göstermektedir. Bu durum, Gemini'nin içerik sunumunda belli ölçüde başarı sağladığını ancak bu başarının genel bir memnuniyet düzeyine ulaşmadığını ortaya koymaktadır.

Kısmen Yeterli Oranının Ağırlığı:

- En yüksek frekans 75 öğrenciyle (yaklaşık %53) "kısmen yeterli" seçeneğindedir. Bu bulgu öğrencilerin büyük bir kısmının Gemini'nin yanıtlarını genel hatlarıyla anlamlı bulduğunu ancak ifadelerin yüzeysel, yetersiz açıklanmış ya da daha sadeleştirilmesi gereken yapıda olduğunu düşündüğünü göstermektedir. Anlaşılırlık açısından tam bir yeterlilik algısı oluşmamıştır.

Olumsuz Görüşler:

- “Yetersiz” diyen 46 öğrenci ile bu oran %32,4’e ulaşmaktadır. Bu oldukça yüksek oran öğrencilerin üçte birinin Gemini’nin sunduğu yanıtları kafa karıştırıcı, yetersiz açıklanmış ya da karmaşık bulduğunu göstermektedir. Bu sonuç Gemini'nin özellikle lise öğrencilerine hitap eden içeriklerde dil ve anlatım düzeyi açısından daha dikkatli olması gerektiğine işaret etmektedir.

Kararsızlar:

- 7 katılımcı (yaklaşık %5) “kararsızım” yanıtını vermiştir. Bu durum öğrencilerin bir kısmının Gemini ile yaşadığı deneyimi net biçimde değerlendirmekte zorlandığını veya çok sınırlı etkileşim yaşadığını düşündürmektedir.

Katılımcıların yaklaşık %63’ü Gemini’nin verdiği cevapları tam ya da kısmen anlaşılır bulmuştur. Ancak bu olumlu görüşlerin ağırlıklı olarak “kısmen yeterli” düzeyinde toplanması ve “yetersiz” değerlendirmesinin %32’yi aşması Gemini’nin anlatım tarzında önemli gelişim alanları olduğunu göstermektedir. Gemini'nin yanıtlarını daha basitleştirilmiş, açıklamalı, örnekli ve öğrenci seviyesine uygun şekilde sunması anlaşılırlık düzeyini artıracaktır. Özellikle lise düzeyindeki fen eğitimi alanında açıklık, yapılandırılmış anlatım ve öğrencinin düzeyine uygun sadeleştirme, anlaşılırlık açısından kritik öneme sahiptir.

Çizelge 4.13.’te öğrencilerin Gemini’ye soru sorma konusundaki değerlendirmeleri gösterilmektedir.

Çizelge 4.13. Gemini’ye Soru Sorma Kolaylığı

Gemini’ye soru sormak kolay ve anlaşılır mıydı?			
Yeterli	Kısmen Yeterli	Kararsızım	Yetersiz
18	75	9	40

Çizelge 4.13.'e göre

Olumlu Değerlendirmeler:

- "Yeterli" ve "Kısmen Yeterli" yanıtları toplamda 18 + 75 = 93 kişi tarafından verilmiştir. Bu da katılımcıların yaklaşık %66'sının Gemini'ye soru sormayı tamamen ya da kısmen kolay ve anlaşılır bulduğunu göstermektedir. Bu çıkarım arayüz veya kullanıcı deneyimi açısından Gemini'nin genel olarak erişilebilir olduğunu fakat bazı eksiklikler barındırdığını da ima eder.

Kısmen Yeterli Oranının Ağırlığı:

- En yüksek yanıt 75 öğrenciyle (yaklaşık %53) "kısmen yeterli" seçeneğidir. Bu bulgu katılımcıların önemli bir kısmının Gemini'ye soru sorma sürecinde kısmen rahat hissettiğini ancak bazı soruların sistem tarafından tam olarak algılanamaması ya da arayüzle ilgili sorunlar nedeniyle sürecin tam anlamıyla sorunsuz ilerlemediğini düşündüğünü gösterir.

Olumsuz Görüşler:

- "Yetersiz" seçeneğini işaretleyen 40 öğrenci (yaklaşık %28) Gemini'nin kullanıcı etkileşimini desteklemede belirgin eksiklikler yaşadığını göstermektedir. Bu durum öğrencilerin Gemini'ye soru yöneltme sürecinde teknik veya dilsel zorluklarla karşılaştığını yönlendirmelerin yetersiz kaldığını ya da yapay zekanın doğru algılama kapasitesinin yeterli bulunmadığını ortaya koymaktadır.

Kararsızlar

- 9 öğrenci (yaklaşık %6) kararsız olduğunu belirtmiştir. Bu durum deneyimlerinin olumlu ya da olumsuz yönde belirgin bir iz bırakmadığını ya da yeterince deneyim yaşamadıklarını göstermektedir.

Katılımcıların yaklaşık %65'lik kısmı Gemini'ye soru sormayı kolay veya kısmen kolay bulmaktadır. Ancak bu değerlendirmelerin büyük kısmı "kısmen yeterli" düzeyinde kalmakta bu da Gemini'nin öğrenciyle etkileşim kurarken daha açıklayıcı yönlendirmelere, sade bir arayüze, dilsel esnekliğe ve daha net komut algılamaya

ihtiyaç duyduğunu göstermektedir. Öte yandan yaklaşık %28'lik "yetersiz" oranı Gemini'nin soru yanıt mekanizmalarında ve kullanıcı etkileşim modellerinde ciddi iyileştirme alanları olduğunu açıkça ortaya koymaktadır. Özellikle lise öğrencileri gibi dijital araçlara aşinalığı farklı düzeylerde olan kullanıcılar için kullanım kolaylığı ve netlik oldukça önemlidir.

Çizelge 4.14.'de öğrencilerin Gemini ile etkileşim deneyim sonuçları gösterilmiştir.

Çizelge 4.14. Gemini ile Etkileşim Deneyimi

Gemini ile etkileşim eğlenceli ve faydalı mıydı?			
Yeterli	Kısmen Yeterli	Kararsızım	Yetersiz
22	59	11	50

Çizelge 4.14.'e göre

Olumlu Değerlendirmeler:

- "Yeterli" ve "Kısmen Yeterli" yanıtları toplamda 22 + 59 = 81 kişi tarafından verilmiştir.
- Bu durum katılımcıların %57'sinin Gemini ile etkileşimi eğlenceli ve faydalı bulduğunu göstermektedir. Bu oran Gemini'nin öğrenciler tarafından kısmi memnuniyetle karşılandığını ve etkileşimin bazı yönlerinin başarılı olduğunu ancak eksikliklerin de mevcut olduğunu düşündürmektedir.

Kısmen Yeterli Oranının Ağırlığı:

- En yüksek yanıt 59 öğrenciyle (yaklaşık %42) "kısmen yeterli" seçeneği olmuştur. Bu durum öğrencilerin büyük bir kısmının Gemini ile etkileşimi tam anlamıyla eğlenceli ve faydalı bulmasa da sürecin öğrenmeye katkı sağladığını ve belirli noktalarda ilgi uyandırdığını düşündüğünü göstermektedir. Yani etkileşim genellikle ilginç ve öğretici olabilmekle birlikte tam anlamıyla tatmin edici olmamıştır.

Olumsuz Görüşler:

- "Yetersiz" yanıtını veren 50 öğrenci (%35,2) Gemini ile etkileşimi eğlenceli veya faydalı bulmadıklarını ifade etmiştir. Bu durum öğrencilerin etkileşim

sürecinde ilgi eksikliği, sıkıcılık veya beklentilerin karşılanmaması gibi olumsuz deneyimler yaşadığını düşündürmektedir. Söz konusu oran etkileşimin önemli bir kısmının öğrenciler açısından sıkıcı veya verimsiz olduğunu göstermektedir.

Kararsızlar

- 11 öğrenci (%7,7) kararsız kalmıştır. Bu durum bazı öğrencilerin Gemini ile etkileşime girdiği halde deneyimlerini net bir şekilde değerlendiremediklerini veya etkileşimin çok kısa ya da belirsiz olduğunu gösteriyor.

Katılımcıların yaklaşık %57'si Gemini ile etkileşimi eğlenceli ve faydalı bulmuş ancak bu değerlendirme çoğunlukla "kısmen yeterli" düzeyinde kalmıştır. Bu durum Gemini'nin öğrenmeye katkı sağladığını ancak etkileşimin her zaman ilgi çekici ve motive edici olmadığını göstermektedir. Öte yandan %35,2'lik "yetersiz" oranı etkileşimin bazı öğrenciler tarafından sıkıcı veya verimsiz bulunduğunu ortaya koymaktadır. Bu bulgular Gemini'nin etkileşim tarzının daha çekici hâle getirilmesi, öğrenci motivasyonunu artıracak şekilde tasarlanması ve öğrenme deneyimini daha eğlenceli kılması gerektiğini göstermektedir.

Anket sonuçları, öğrencilerin Gemini'yi tekrar kullanma eğilimlerinin düşük olduğunu göstermektedir. Öğrenciler, Gemini'nin doğruluk ve açıklayıcılık düzeylerini yetersiz bulmuş; bu durum, aracın yeniden kullanılma isteğini olumsuz yönde etkilemiştir. Çizelge 4.15.'te öğrencilerin Gemini'yi tekrar kullanma eğilim sonuçları verilmiştir.

Çizelge 4.15. Gemini Tekrar Kullanım Eğilimi

Gemini'yi tekrar kullanmayı düşünür müsünüz?		
Evet	Hayır	Kararsızım
49	38	55

Çizelge 4.15.'e göre

Araştırmaya katılan 142 öğrenciden elde edilen anket verileri aşağıdaki gibidir.

Olumlu Yaklaşım:

- “Evet” diyen 49 kişi yani katılımcıların yaklaşık %35’i Gemini’yi tekrar kullanmayı düşündüğünü belirtmiştir. Bu oran Gemini’nin bazı öğrenciler için faydalı ve tekrar tercih edilebilir olduğunu göstermektedir ancak çoğunluk değildir.

Kararsızların Fazlalığı:

- 55 öğrenci (yaklaşık %39) “kararsızım” yanıtını vermiştir. Bu oran oldukça yüksektir ve birçok öğrencinin Gemini hakkında net bir görüş oluşturamadığını belki de daha fazla deneyim veya iyileştirme beklediğini göstermektedir. Bu grup, doğru yönlendirmeler ve geliştirilen içeriklerle yeniden kazanılabilir potansiyele sahiptir.

Olumsuz Görüş:

- “Hayır” diyen 38 kişi (%26,8) Gemini’yi tekrar kullanmak istemediğini belirtmiştir. Bu oran azımsanamayacak düzeydedir ve Gemini’nin içerik kalitesi, kullanım kolaylığı veya ilgi çekiciliği konularında bazı yetersizliklerin olduğunu düşündürebilir.

Veriler Gemini’nin tekrar kullanım niyetinde ChatGPT’ye göre daha düşük düzeyde tercih edildiğini ortaya koymaktadır. Öğrencilerin %34,5’i Gemini kullanımına olumlu yaklaşırken %38,7’si kararsız, %26,8’i ise olumsuz bir tutum sergilemektedir. Bu durum Gemini’nin öğrenci beklentilerini tam olarak karşılamadığına özellikle açıklık, doğruluk ya da etkileşim yönlerinden geliştirilmesi gerektiğine işaret etmektedir. Kararsız grubun fazlalığı Gemini için potansiyel iyileştirme alanlarının olduğunu ve doğru müdahalelerle kullanıcı memnuniyetinin artırılabileceğini göstermektedir.

Katılımcılara “Gemini'nin Daha İyi olmasını Sağlayacak Bir Öneriniz Var Mı?” sorusu opsiyonel olacak şekilde yöneltilmiştir. Sonuçlar lise öğrencilerinin üretken yapay zekâ araçlarından Gemini’ye yönelik deneyimlerini ve geri bildirimlerini içermektedir. Öğrencilere Gemini'nin kullanım sürecinde karşılaştıkları olumlu ve olumsuz durumları içeren açık uçlu sorular yöneltilmiş, yanıtlar tematik analiz yöntemiyle çözümlenmiştir.

- Görsel ve Çizim Yetersizliği

En sık tekrar eden şikayetlerden biri Gemini'nin şekil ve görsel çizim yeteneğiyle ilgilidir: "Gemini çizim yapamıyor", "çizimi yanlış yapıyor", "görseller eksik", "resim çizemedi" gibi ifadeler çok sayıda öğrencinin yorumlarında yer almaktadır. Öğrenciler özellikle "resim çiz" gibi talepler karşısında Gemini'nin beklentileri karşılamadığını ve görsellerin ya hiç oluşturulmadığını ya da yanlış oluşturulduğunu vurgulamışlardır.

Öneri: Görsel içerik üretme becerilerinin artırılması ve çizim araçlarının daha güvenilir hale getirilmesi önemlidir.

- Cevapların Uzunluğu ve Aşırı Akademik Dili

"Cevaplar çok uzun", "gereksiz bilgiler veriyor", "akademik seviyede", "konu dışına çıkıyor" gibi eleştiriler oldukça yaygındır. Özellikle lise düzeyindeki kullanıcılar cevapların seviyelerine uygun olmadığını belirtmiştir.

Öneri: Yanıtlar sadeleştirilmeli, hedef kitleye uygun dil, uygun örnekler kullanılmalıdır.

- Yanlış veya Tutarsız Cevaplar

"Yanlış cevaplar veriyor", "ChatGPT'ye göre daha çok hata yapıyor", "iki şık da yanlış" gibi yorumlar Gemini'nin doğruluk oranı konusunda güven sorunu yaşandığını göstermektedir. ChatGPT ile karşılaştırıldığında daha çok hata yaptığı sıklıkla dile getirilmiştir.

Öneri: Akademik doğruluk oranının artırılması ve doğrulama algoritmalarının geliştirilmesi gerekmektedir.

- Karşılaştırmalı Olumsuz Değerlendirmeler

Çok sayıda öğrenci Gemini'yi doğrudan ChatGPT ile karşılaştırmış ve ChatGPT'yi daha kapsamlı, sade, doğru ve güvenilir bulduğunu ifade etmiştir. "ChatGPT daha iyi",

“Gemini’nin yanına yaklaşamaz”, “asla Gemini kullanmam” gibi yorumlar bu görüşü desteklemektedir.

Öneri: Kullanıcı deneyimi açısından Gemini’nin yapay zekâ rekabetinde daha güçlü hale gelmesi için kullanıcı dostu arayüz ve kaliteli içerik sunumu öncelikli hedef olmalıdır.

- Genel Değerlendirme ve Geliştirme Önerileri

Bazı öğrenciler doğrudan geliştirme önerileri sunmuştur:

- “Cevaplar kısa ve öz olmalı.”
- “Görseller geliştirilmeli.”
- “ChatGPT’ye satılsın.”
- “Google’dan alıntı yapmasın, kendi üretmeli.”
- “Cevaplarda duygu/düşünce de olsun.”

Öğrenciler Gemini’nin bazı yönlerinin (örneğin cevap uzunluğu, çizim yetkinliği ve doğruluk) geliştirilmesi gerektiğini belirgin bir şekilde vurgulamıştır. Kullanıcılar Gemini’nin bilgi sunumundaki yoğunluk ve akademik anlatım sebebiyle erişilebilirlikte zorlandıklarını belirtmektedir. ChatGPT ile kıyaslandığında Gemini şu anda öğrenciler için ikinci tercih konumundadır.

4.2.1. Farklı Okul Türlerine Göre Gemini Ortalama Puanlarının Karşılaştırılması (Kruskal-Wallis Testi Sonuçları)

```
NPAR TESTS
        /K-W=Gemini_Ortalama BY OkulAdı (5 6 7)
        /MISSING ANALYSIS.
```

NPar Tests
Kruskal-Wallis Test

Ranks

Okul Adı		N	Mean Rank
Gemini_Ortalama	ŞAİK AL	50	47,33
	Gazi Sosyal Bilimler Lisesi	42	96,82
	Fen Lisesi	50	74,40
	Total	142	

Test Statistics[a,b]

	Gemini_Ortalama
Chi-Square	33,799
Df	2
Asymp. Sig.	0,000

a. Kruskal-Wallis Test
b. Grouping Variable: Okul Adı

Şekil 4.2. Okul Bazında Gemini Ortalama Puanları ve Kruskal-Wallis Sonuçları

Şekil 4.2.'de öğrencilerin değerlendirmesine göre okul bazında Gemini için ortalama puanlar ve Kruskal-Wallis test sonuçları verilmiştir. Kruskal-Wallis testi, üç veya daha fazla grubun medyanlarının birbirinden farklı olup olmadığını test etmek için kullanılan parametrik olmayan bir testtir. Şekilde "Gemini_Ortalama" isimli bir değişkenin farklı okullara (ŞAİK AL, Gazi Sosyal Bilimler Lisesi, Fen Lisesi) göre farklılık gösterip göstermediği test edilmiştir.

4.2.1.1. Tanımlayıcı İstatistikler

- Gemini_Ortalama: Öğrencilerin Gemini kullanımına yönelik ortalama puanları 33,79 düzeyindedir.
- Okul Adı: Sayısal kodlarla temsil edilen üç okul türü arasında dağılım yapıldı. (5 = ŞAİK AL, 6 = Gazi Sosyal Bilimler Lisesi, 7 = Fen Lisesi)

4.2.1.2. Kruskal-Wallis Mean Rank (Sıra Ortalamaları)

Çizelge 4.16.'da Gemini'nin okullara göre ortalama sıra dağılım sonuçları gösterilmiştir.

Çizelge 4.16. Gemini'nin Okullara Göre Ortalama Sıra Dağılımı

Okul Adı	N	Ortalama Sıra (Mean Rank)
ŞAİK AL	50	47,33
Gazi Sosyal Bilimler Lisesi	42	96,82
Fen Lisesi	50	74,40
Toplam	142	

Gazi Sosyal Bilimler Lisesi en yüksek ortalama sıraya sahiptir (96,82). Bu durum bu okulda "Gemini_Ortalama" değerlerinin daha yüksek olduğunu gösterir. En düşük ortalama sıra ise ŞAİK AL'ye ait (47,33), bu durum bu okulun "Gemini_Ortalama" açısından diğerlerinden düşük değerlere sahip olduğunu gösterir.

4.2.1.3. Kruskal-Wallis Test İstatistikleri

Çizelge 4.17.'de okullar arası Gemini ortalama puan farklılıkları gösterilmiştir.

Çizelge 4.17. Okullar Arası Gemini Ortalama Puan Farklılıkları

Test	Değer
Chi-Square	33,799
df (Serbestlik derecesi)	2
Asymp. Sig. (p değeri)	0,000

Çizelge 4.17.'a göre

- Anlamlılık düzeyi (p): 0,000 → bu, $p < 0.05$ anlamına gelir.
- Kruskal-Wallis testi sonucunda istatistiksel olarak anlamlı bir fark bulunmuştur ($p = 0,000 < 0.05$).

Farklı okullara göre "Gemini_Ortalama" değişkeninde anlamlı farklılıklar vardır. Gazi Sosyal Bilimler Lisesi ortalamada en yüksek sıraya sahiptir, ŞAİK AL ise en düşük sıradadır.

4.2.1.4. Genel Sonuç ve Yorum

ŞAİK Anadolu Lisesi öğrencileri en düşük ortalama sıralamaya sahiptir. Bu da onların Gemini'den daha az verim aldığı veya başarılarının daha düşük olduğu anlamına gelebilir. Farklı okul türleri arasında Gemini_Ortalama değişkeni açısından istatistiksel olarak anlamlı fark vardır. Öğrencilerin okudukları lise türü Gemini'den elde ettikleri başarıyı etkilemektedir.

Çizelge 4.18.'de Öğrenciler tarafından yapılan ChatGPT ve Gemini'nin okullar arası performans ve tercih karşılaştırılması verilmiştir. Çizelge 4.18.'de görüldüğü gibi ChatGPT ve Gemini üretken yapay zekâ araçlarının lise öğrencileri tarafından değerlendirilmesi sonucunda kimya ve biyoloji derslerinin eğitiminde ChatGPT'nin kullanımı Gemini'ye kıyasla öğrenciler tarafından daha çok tercih edilmektedir.

Çizelge 4.18. ChatGPT ve Gemininin Okullar Arası Performans Karşılaştırması

Değişken	**ChatGPT**	**Gemini**
Katılımcı Sayısı (N)	142	142
Okul Türleri	ŞAİK AL Gazi SBL Fen Lisesi	ŞAİK AL Gazi SBL Fen Lisesi
Ortalama Sıra (Mean Rank)	85,10 64,05 64,16	47,33 96,82 74,40
Test İstatistiği (χ^2)	8,559	33,799
Serbestlik Derecesi (df)	2	2
Anlamlılık Düzeyi (p)	0,014	0,000
Yorum	Okullar arasında anlamlı fark var	Okullar arasında anlamlı fark var
Tekrar Kullanma Eğilimi (%)	%90,8 Evet %0,7 Hayır %8,5 Kararsız	%34,5 Evet %26,8 Hayır %38,7 Kararsız
Genel Tercih	Yüksek	Düşük

4.3. Ders Bazlı Değerlendirmeler: Biyoloji ve Kimya

Biyoloji ve Kimya derslerinde ChatGPT ve Gemini ortalamalarının okul türüne göre farklılaşıp farklılaşmadığını Kruskal-Wallis testiyle incelenmiştir.

4.3.1. Farklı Lise Türlerinde Öğrencilerin ChatGPT'nin Biyoloji Dersindeki Kullanımına İlişkin Değerlendirmelerinin Karşılaştırılması

Şekil 4.3.'de ChatGPT'nin biyoloji dersinde kullanımı lise türlerine göre karşılaştırılmış ve sonuçları gösterilmiştir.

```
USE ALL .
COMPUTE filter_$=(DersAdı=1) .
VARIABLE LABELS filter_$ 'DersAdı=1 (FILTER)' .
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected' .
FORMATS filter_$ (f1.0) .
FILTER BY filter_$ .
EXECUTE .
NPAR TESTS
        /K-W=ChatGPT_Ortalama BY OkulAdı (5 7)
        /STATISTICS DESCRIPTIVES
        /MISSING ANALYSIS.
```

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
ChatGPT_Ortalama	75	2,2822	0,40638	1,33	2,83
Okul Adı	75	6,00	0,838	5	7

Kruskal-Wallis Test

Ranks

Okul Adı		N	Mean Rank
ChatGPT_Ortalama	ŞAİK AL	26	42,94
	Gazi Sosyal Bilimler Lisesi	23	41,09
	Fen Lisesi	26	30,33
	Total	75	

Test Statistics[a,b]

	ChatGPT_Ortalama
Chi-Square	5,153
Df	2
Asymp. Sig.	0,076

a. Kruskal-Wallis Test
b. Grouping Variable: Okul Adı

Şekil 4.3. ChatGPT'nin Biyoloji Dersinde Kullanımı: Lise Türlerine Göre Karşılaştırma

Analiz Kapsamı:

- Ders: Biyoloji (DersAdı = 1)
- Amaç: Farklı lise türlerindeki öğrencilerin ChatGPT'nin biyoloji dersindeki kullanımına dair değerlendirmeleri arasında anlamlı bir fark olup olmadığına bakılmıştır.
- Bu çalışma kapsamında öğrencilerin biyoloji dersinde ChatGPT kullanımına ilişkin değerlendirmeleri okul türüne göre karşılaştırılmıştır. Yapılan Kruskal-Wallis testi sonucunda okul türleri arasında istatistiksel olarak anlamlı bir fark bulunmamıştır (Chi-Square = 5,153 ; p = ,076).
- Ancak sıralı ortalama puanlara bakıldığında Şehit Ali İhsan Kalmaz Anadolu Lisesi (42,94) ve Gazi Sosyal Bilimler Lisesi (41,09) öğrencileri ChatGPT'yi daha yüksek düzeyde olumlu değerlendirmiştir. Fen Lisesi öğrencilerinin ortalama sıralaması ise daha düşüktür (30,33).

Bu bulgular farklı lise türlerindeki öğrencilerin ChatGPT'nin biyoloji dersindeki faydasına ilişkin algılarının farklılaştığını göstermektedir. Özellikle ŞAİK Anadolu Lisesi ve Sosyal Bilimler Lisesi öğrencileri ChatGPT'nin biyoloji dersindeki katkısını daha olumlu değerlendirmiştir. Bu durum fen lisesi öğrencilerinin konuya daha eleştirel yaklaştığını veya beklentilerinin daha yüksek olduğunu düşündürebilir.

4.3.2. Farklı Lise Türlerinde Öğrencilerin ChatGPT'nin Kimya Dersindeki Kullanımına İlişkin Değerlendirmelerinin Karşılaştırılması

Şekil 4.4.'te ChatGPT'nin kimya dersinde kullanımının karşılaştırılması lise türlerine göre gösterilmiştir.

```
USE ALL .
COMPUTE filter_$=(DersAdı=2) .
VARIABLE LABELS filter_$ 'DersAdı=2 (FILTER)' .
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected' .
FORMATS filter_$ (f1.0) .
FILTER BY filter_$ .
EXECUTE .
NPAR TESTS
        /K-W=ChatGPT_Ortalama BY OkulAdı (5 7)
        /STATISTICS DESCRIPTIVES
        /MISSING ANALYSIS.
```

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
ChatGPT_Ortalama	67	2,0547	0,46907	0,67	2,83
Okul Adı	67	6,00	0,853	5	7

Kruskal-Wallis Test

Ranks

	Okul Adı	N	Mean Rank
ChatGPT_Ortalama	ŞAİK AL	24	42,77
	Gazi Sosyal Bilimler Lisesi	19	22,68
	Fen Lisesi	24	34,19
	Total	67	

Test Statistics[a,b]

	ChatGPT_Ortalama
Chi-Square	11,432
Df	2
Asymp. Sig.	0,003

a. Kruskal-Wallis Test
b. Grouping Variable: Okul Adı

Şekil 4.4. ChatGPT'nin Kimya Dersinde Kullanımı: Lise Türlerine Göre Karşılaştırma

Analizin Kapsamı:

- Ders: Kimya (DersAdı = 2)
- ChatGPT_Ortalama; Öğrencilerin ChatGPT'ye ilişkin ortalama değerlendirmeleri
- Okul türü: ŞAİK AL, Gazi Sosyal Bilimler Lisesi, Fen Lisesi

Çizelge 4.19.'da Kimya Dersine ait ChatGPT'nin ortalama değerlendirmeleri sunulmuştur.

Çizelge 4.19. Kimya Dersinde ChatGPT_Ortalama Değerlendirmeleri

Değişken	N	Ortalama	Std. Sapma	Min	Max
ChatGPT_Ortalama	67	2,0547	0,46907	0,67	2,83

Çizelge 4.19.'a göre

- Öğrencilerin ChatGPT'ye ilişkin değerlendirme ortalaması: 2,05
- Bu ortalama Kimya dersinde 2,05; Biyoloji dersine göre daha düşük (Biyoloji'de 2,28).

Çizelge 4.20'de kimya dersi için yapılan ChatGPT' nin değerlendirmelerinde okullar arası farklılık gösterilmiştir.

Çizelge 4.20. ChatGPT'nin Kimya Dersindeki Değerlendirmeleri: Okullar Arası Farklılık

Okul Adı	N	Ortalama Sıra
ŞAİK AL	24	42,77
Gazi Sosyal Bilimler Lisesi	19	22,68
Fen Lisesi	24	34,19

Çizelge 4.20.'ye göre

- ŞAİK AL öğrencileri ChatGPT'yi kimya dersinde en olumlu değerlendiren grup.
- Gazi Sosyal Bilimler Lisesi öğrencileri en düşük değerlendirmeyi yapmış.
- Fen Lisesi öğrencileri orta düzeyde değerlendirme yapmış.

Çizelge 4.21'de genel istatistiksel sonuçlar verilmiştir.

Çizelge 4.21. İstatistiksel Sonuçlar

Test	Değer
Chi-Square	11,432
df	2
Asymp. Sig.	0,003

Çizelge 4.21.'e göre

- p-değeri: 0,003 <0,05 → Fark istatistiksel olarak anlamlı.
- Yani okul türüne göre ChatGPT'nin kimya dersindeki değerlendirmelerinde anlamlı farklılık vardır.

Yapılan Kruskal-Wallis testi sonucuna göre öğrencilerin ChatGPT'nin kimya dersindeki kullanımına ilişkin değerlendirmeleri okul türüne göre anlamlı düzeyde farklılık göstermektedir (Chi-Square = 11,432, p = 0,003).

Ortalama sıralamalara göre Şehit Ali İhsan Kalmaz Anadolu Lisesi öğrencileri ChatGPT'yi kimya dersinde en olumlu değerlendiren grup olmuştur (Mean Rank = 42,77). Fen Lisesi öğrencileri orta düzeyde değerlendirme yaparken (34,19) Gazi Sosyal Bilimler Lisesi öğrencileri en düşük ortalamayı vermiştir (22,68).

Bu bulgular ChatGPT'nin kimya dersindeki etkisinin okul türüne göre farklı algılandığını ve bazı okul türlerinde daha fazla fayda sağladığını düşündürmektedir. Sosyal Bilimler Lisesi öğrencilerinin düşük puan vermesi bu öğrencilerin fen ağırlıklı derslerdeki teknolojik araçlara olan ilgisinin veya yeterliliğinin daha sınırlı olabileceğini düşündürebilir.

4.3.3. Farklı Lise Türlerinde Öğrencilerin Gemini'nin Biyoloji Dersindeki Kullanımına İlişkin Değerlendirmelerinin Karşılaştırılması

```
USE ALL .
COMPUTE filter_$=(DersAdı=1) .
VARIABLE LABELS filter_$ 'DersAdı=1 (FILTER)' .
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected' .
FORMATS filter_$ (f1.0) .
FILTER BY filter_$ .
EXECUTE .
NPAR TESTS
        /K-W=Gemini_Ortalama BY OkulAdı (5 7)
        /MISSING ANALYSIS.
```

NPar Tests

Kruskal-Wallis Test

Ranks

	Okul Adı	N	Mean Rank
Gemini_Ortalama	ŞAİK AL	26	24,42
	Gazi Sosyal Bilimler Lisesi	23	52,89
	Fen Lisesi	26	38,40
	Total	75	

Test Statistics[a,b]

	Gemini_Ortalama
Chi-Square	21,052
Df	2
Asymp. Sig.	0,000

a. Kruskal-Wallis Test
b. Grouping Variable: Okul Adı

Şekil 4.5. Gemini'nin Biyoloji Dersinde Kullanımı: Lise Türlerine Göre Karşılaştırma

Çizelge 4.18.'de görüldüğü gibi ChatGPT ve Gemini üretken yapay zekâ araçlarının lise öğrencileri tarafından değerlendirilmesi sonucunda kimya ve biyoloji derslerinin eğitiminde ChatGPT'nin kullanımı Gemini'ye kıyasla öğrenciler tarafından daha çok tercih edilmektedir.

Bu analiz “Gemini_Ortalama” değişkeninin sadece “DersAdı = Biyoloji” olan öğrencilerde farklı okul türlerine (ŞAİK AL, Gazi Sosyal Bilimler Lisesi, Fen Lisesi) göre farklılık gösterip göstermediğini incelemek amacıyla Kruskal-Wallis testi ile gerçekleştirilmiştir.

Elde edilen bulgulara göre:

- ŞAİK AL için ortalama sıra değeri: 24,42
- Gazi Sosyal Bilimler Lisesi için ortalama sıra değeri: 52,89
- Fen Lisesi için ortalama sıra değeri: 38,40

Test istatistikleri incelendiğinde Ki-Kare değeri 21,052, serbestlik derecesi (df) 2 ve anlamlılık değeri (p) 0,000 olarak bulunmuştur ($p < 0,05$). Bu sonuç en az iki okulun “Gemini_Ortalama” değerleri arasında istatistiksel olarak anlamlı bir farklılık olduğunu göstermektedir.

Sonuç olarak Biyoloji dersini seçmiş öğrenciler arasında okullara göre anlamlı bir farklılık vardır. En yüksek ortalama sıraya sahip okul Gazi Sosyal Bilimler Lisesi, en düşük sıraya sahip okul ise ŞAİK AL'dir.

4.3.4. Farklı Lise Türlerinde Öğrencilerin Gemini’nin Kimya Dersindeki Kullanımına İlişkin Değerlendirmelerinin Karşılaştırılması

Şekil 4.6.’da Gemini’nin kimya dersinde kullanımına ilişkin değerlendirmeler lise türlerine göre karşılaştırmalı olarak gösterilmiştir.

```
USE ALL .
COMPUTE filter_$=(DersAdı=2) .
VARIABLE LABELS filter_$ 'DersAdı=2 (FILTER)' .
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected' .
FORMATS filter_$ (f1.0) .
FILTER BY filter_$ .
EXECUTE .
NPAR TESTS
        /K-W=Gemini_Ortalama BY OkulAdı (5 7)
        /MISSING ANALYSIS.
```

Npar Tests
Kruskal-Wallis Test

Ranks

Okul Adı		N	Mean Rank
Gemini_Ortalama	ŞAİK AL	24	23,23
	Gazi Sosyal Bilimler Lisesi	19	44,34
	Fen Lisesi	24	36,58
	Total	67	

Test Statistics[a,b]

	Gemini_Ortalama
Chi-Square	13,286
Df	2
Asymp. Sig.	0,001

a. Kruskal-Wallis Test
b. Grouping Variable: Okul Adı

Şekil 4.6. Gemini'nin Kimya Dersinde Kullanımı: Lise Türlerine Göre Karşılaştırma

Bu analiz "Gemini_Ortalama" değişkeninin sadece "DersAdı = Kimya" olan öğrencilerde farklı okul türlerine (ŞAİK AL, Gazi Sosyal Bilimler Lisesi, Fen Lisesi) göre farklılık gösterip göstermediğini incelemek amacıyla Kruskal-Wallis testi ile gerçekleştirilmiştir.

Elde edilen bulgulara göre:

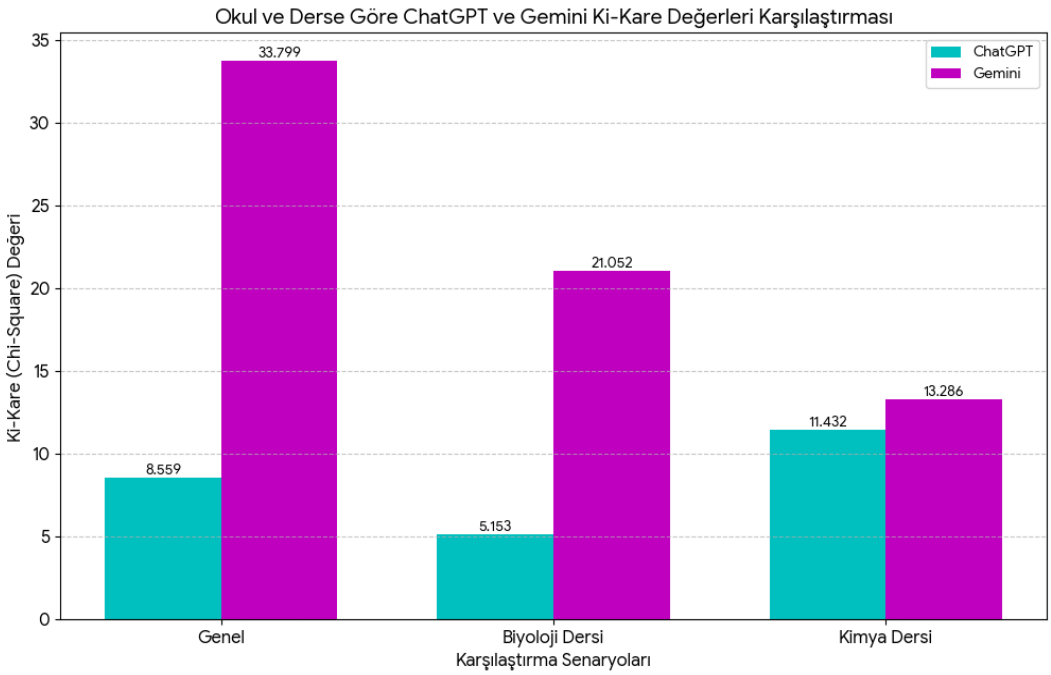
- ŞAİK AL için ortalama sıra değeri: 23,23

- Gazi Sosyal Bilimler Lisesi için ortalama sıra değeri: 44,34
- Fen Lisesi için ortalama sıra değeri: 36,58

Test istatistikleri incelendiğinde Ki-Kare değeri 13,286, serbestlik derecesi (df) 2 ve anlamlılık değeri (p) 0,001 olarak bulunmuştur ($p < 0,05$). Bu sonuç en az iki okulun "Gemini_Ortalama" değerleri arasında istatistiksel olarak anlamlı bir farklılık olduğunu göstermektedir.

Sonuç olarak Kimya dersini seçmiş öğrenciler arasında okullara göre anlamlı bir farklılık vardır. En yüksek ortalama sıraya sahip okul Gazi Sosyal Bilimler Lisesi, en düşük sıraya sahip okul ise ŞAİK AL'dir.

Şekil 4.7.'de okul ve derse göre ChatGPT ile Gemini'nin kullanımına ilişkin Ki-Kare değerleri karşılaştırmalı olarak sunulmuştur.



Şekil 4.7. Okul ve derse göre ChatGPT ve Gemini Ki-Kare Değerleri Karşılaştırması

Şekil 4.7.'de Kruskal-Wallis testine ait Ki-Kare değerlerini göstermektedir. Bu değerler karşılaştırılan gruplar arasındaki farkın büyüklüğünü belirtir. Değer ne kadar yüksekse gruplar arasındaki farkın o kadar büyük olduğu anlamına gelir.

5. SONUÇ VE ÖNERİLER

5.1. Sonuçlar

Bu araştırmada üretken yapay zekâ araçları olan ChatGPT ve Gemini'nin lise düzeyindeki kimya ve biyoloji derslerinde öğrencilerin öğrenme süreçlerine etkisi incelenmiştir. Uygulama üç farklı lisede gerçekleştirilmiş ve hem nicel hem de nitel verilerle desteklenmiştir.

Elde edilen bulgular ChatGPT'nin, lise düzeyindeki öğrenciler tarafından Gemini'ye kıyasla çok daha işlevsel, açıklayıcı ve yeniden kullanılabilir bir araç olarak değerlendirildiğini göstermektedir. Katılımcıların büyük çoğunluğu ChatGPT'nin sunduğu yanıtları doğruluk, anlaşılırlık ve açıklayıcılık açısından yeterli bulmuş ve özellikle kavramsal açıklamalarda bilgi sunumunun sadeliği ve doğruluğu öğrencilerin beklentilerini yüksek oranda karşılamıştır. Buna karşın Gemini, görsel üretim, cevap doğruluğu, dil seviyesi ve kullanıcı etkileşimi açısından daha sınırlı bir performans sergilemiş ve öğrencilerde yeterli güven duygusu oluşturamamıştır.

ChatGPT ile Gemini arasındaki en belirgin fark öğrenci tatmini ve yeniden kullanım isteğidir. ChatGPT daha fazla öğrenciye hitap eden daha anlaşılır ve pedagojik olarak daha etkili bir model olarak öne çıkmaktadır. Ayrıca öğrenciler ChatGPT'nin verdiği yanıtları daha sade, eğlenceli ve anlaşılır bulurken Gemini'yi daha akademik, daha karmaşık ve zaman zaman hatalı olarak tanımlanmıştır.

Araştırma sonuçları genel olarak ChatGPT'nin lise düzeyindeki öğrenciler arasında her okul türünde benzer biçimde olumlu değerlendirildiğini, buna karşın Gemini'ye yönelik memnuniyet düzeylerinin okul türüne göre anlamlı farklılıklar gösterdiğini ortaya koymuştur. ChatGPT'nin özellikle kimya dersinde öğrencilerin kavramsal anlamalarını kolaylaştırdığı ve derslere olan ilgiyi artırdığı belirlenmiştir. Açık uçlu öğrenci görüşleri de bu bulguları destekler nitelikte olup, ChatGPT'nin anlaşılır, yönlendirici ve tekrar edilebilir yanıtlar sunduğunu göstermektedir.

Buna karşın Gemini, hem kimya hem biyoloji derslerinde sınırlı etki göstermiş, bazı durumlarda yanıtsız veya yetersiz açıklamalar sunmuştur. Her iki araç da öğrencilerin

derse ilgisini artırmış olsa da ChatGPT, öğrenciyle etkileşim sürecinde daha "kapsayıcı" ve "öğrenci dilinde yanıt verme" yeteneğiyle öne çıkmıştır.

İstatistiksel analizler (Kruskal–Wallis testi) Anadolu Lisesi öğrencilerinin ChatGPT'ye yönelik değerlendirmelerinin diğer okul türlerine kıyasla daha yüksek olduğunu, Fen ve Sosyal Bilimler Lisesi öğrencilerinin ise daha temkinli bir tutum sergilediğini göstermiştir. Gemini için ise tüm okul türlerinde düşük memnuniyet düzeyi ortak bir sonuç olarak belirlenmiştir.

Sonuç olarak bu araştırma üretken yapay zekâ araçlarının fen bilimleri eğitimi için ciddi potansiyel barındırdığını ancak araçların doğruluk, dil sadeliği, görselleştirme yetenekleri ve kullanıcı deneyimi gibi alanlarda farklı performanslar sergilediğini göstermiştir. ChatGPT şu anki haliyle fen dersleri için daha güçlü bir alternatif sunarken Gemini'nin lise düzeyindeki eğitsel kullanımı açısından belirli iyileştirmelere ihtiyaç duyduğu açıktır. Bu sonuçlar, üretken yapay zekâ teknolojilerinin eğitim ortamlarında kullanılabilirliğine ilişkin önemli ipuçları sunmakta özellikle ChatGPT benzeri dil modellerinin pedagojik süreçlerde daha etkili biçimde bütünleştirilebileceğini ortaya koymaktadır.

5.2. Öneriler

Araştırmadan elde edilen sonuçlara dayanarak aşağıdaki öneriler geliştirilmiştir:

- ChatGPT ve Gemini'nin etkililiği farklı yaş gruplarında, farklı ders düzeylerinde ve branşlarda test edilmelidir. Bu çalışma yalnızca 10. sınıftaki öğrencilerle ve fen dersleriyle sınırlı olduğundan farklı bağlamlarda yapılacak çalışmalar genellenebilirliği artıracaktır.
- ChatGPT ile Gemini'nin gelişen sürümleriyle uzun dönemli etki analizleri yapılmalıdır. Mevcut veriler kısa dönemli deneyimlere dayalıdır, zamanla değişen tutum ve öğrenme düzeyleri için takip çalışmaları gereklidir.
- Yapay zekâ araçlarının öğretmen tutumları ve öğretim sürecine etkisi ayrı bir araştırma konusu olarak ele alınmalıdır. Öğretmenlerin teknolojiye yönelik algısı sınıf içi uygulamaları doğrudan etkilemektedir.

- Yapay zekâ araçlarının veri güvenliği, etik kullanımı ile içerik doğruluğu gibi yönleri de derinlemesine araştırılmalıdır. Özellikle öğrenci verilerinin gizliliği ve doğru bilgiye erişim, eğitimde yapay zekânın temel sorun alanları arasında yer almaktadır.
- Üretken yapay zekânın uzun vadede öğrencilerin akademik başarıları, öğrenme kalıcılığı ve eleştirel düşünme becerilerine etkisi incelenmelidir.
- ChatGPT ve Gemini dışında diğer büyük dil modellerinin de dahil edildiği karşılaştırmalı çalışmalar gerçekleştirilmelidir. Böylece eğitimde farklı yapay zekâ yaklaşımlarının pedagojik katkıları daha bütüncül biçimde değerlendirilebilir.

KAYNAKLAR

Agbong-Coates, I. J. G. (2024). ChatGPT integration significantly boosts personalized learning outcomes: A Philippine study. International Journal of Educational Management and Development Studies, 5(2), 165–186.

Aktay, S., Gök, S., & Uzunoğlu, D. (2023). ChatGPT in education. Türk Akademik Yayınlar Dergisi (TAY Journal), 7(2), 378–406.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., … Amodei, D. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877–1901.

Dao, X.-Q., Le, N.-B., Vo, T.-D., Ngo, B.-B., & Phan, X.-D. (2023). LLMs' capabilities at the high school level in chemistry: Cases of ChatGPT and Microsoft Bing Chat. ChemRxiv Preprint. https://chemrxiv.org/engage/chemrxiv/article-details/648d16b2e64f843f41e5d6bd

de Zárate, D. O., Castañeda, R., Mercade, L., Gómez, V. J., Rubio, A. D., Díaz-Fernández, F. J., & Pinilla-Cienfuegos, E. (2024). Enhancing K-12 students' performance in chemistry through ChatGPT-powered blended learning in the Education 4.0 era. In INTED2024 Proceedings (3920–3929). IATED.

Du-Harpur, X., Watt, F. M., Luscombe, N. M., & Lynch, M. D. (2020). What is AI? Applications of artificial intelligence to dermatology. British Journal of Dermatology, 183(3), 423–430.

Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2024). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots' impact on university teachers' assessment practices. Assessment & Evaluation in Higher Education, 49(3), 363–375.

Farquhar, S., Kossen, J., Kuhn, L., & Gal, Y. (2024). Detecting hallucinations in large language models using semantic entropy. Nature, 630(8017), 625–630.

Firdaus, T., Sholeha, S. A., Jannah, M., & Setiawan, A. R. (2024). Comparison of ChatGPT and Gemini AI in answering higher-order thinking skill biology questions: Accuracy and evaluation. International Journal of Science Education and Teaching, 3(3), 126–138.

Fryer, L. K., & Dinsmore, D. L. (2023). What does current GenAI actually mean for student learning? ResearchGate.

Gan, W., Qi, Z., Wu, J., & Lin, J. C. W. (2023b, December). Large language models in education: Vision and opportunities. In 2023 IEEE International Conference on Big Data (BigData) (4776–4785). IEEE.

Gan, W., Ye, Z., Wan, S., & Yu, P. S. (2023a, April). Web 3.0: The future of internet. In Companion Proceedings of the ACM Web Conference 2023 (1266–1275). ACM.

Huang, Y., Zhang, Q., & Lin, J. (2022). The impact of AI-based learning assistants on student engagement in science classes. Journal of Educational Technology, 45(2), 112–128.

Imran, M., & Almusharraf, N. (2024). Google Gemini as a next generation AI educational tool: A review of emerging educational technology. Smart Learning Environments, 11(1), 22.

Kamruzzaman, M. (2023). AI- and IoT-assisted sustainable education systems during pandemics, such as COVID-19, for smart cities. Sustainability, 15(10), 8354.

Karabıyık, Ü. (2024). Matematik eğitiminde yenilikçi bir yaklaşım: ChatGPT'nin rolü. Uşak Üniversitesi Eğitim Araştırmaları Dergisi, 10(1), 26–46.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274.

Katiyar, P. (2024). AI-driven personalized learning systems: Enhancing educational effectiveness. Educational Administration: Theory and Practice, 30(5), 11514–11524.

Kharchenko, Y. V., & Babenko, O. M. (2024, September). Advantages and limitations of large language models in chemistry education: A comparative analysis of ChatGPT, Gemini and Copilot. In Proceedings of the Free Open-Access Proceedings for Computer Science Workshops, 42–59.

Kim, T. W. (2023). Application of artificial intelligence chatbot, including ChatGPT, in education: Scholarly work, programming, and content generation, and its prospects: A narrative review. Journal of Educational Evaluation for Health Professions, 20, 14.

Leite, B. S. (2023). Artificial intelligence and chemistry teaching: A propaedeutic analysis of ChatGPT in chemical concepts defining. Química Nova, 46(9), 915–923.

Monib, W. K. (2024). Generative AI and future education: A review, theoretical validation, and authors' perspective on challenges and solutions. PeerJ Computer Science, 10, e2105.

Nguyen, T., Cao, L., Nguyễn, P., Tran, V., & Nguyễn, P. (2023). Capabilities, benefits, and role of ChatGPT in chemistry teaching and learning in Vietnamese high schools. EdArXiv Preprint. https://doi.org/10.35542/osf.io/4wt6q

Nurjanah, A. (2024). Artificial intelligence (AI) usage in today's teaching and learning process: A review. Syntax Idea, 6(3), 1517–1523.

Rane, N., Choudhary, S., & Rane, J. (2024). Gemini versus ChatGPT: Applications, performance, architecture, capabilities, and implementation. Journal of Applied Artificial Intelligence, 5(1), 69–93.

Russell, S. J., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.

Topal, D. A., Dilek Eren, C., & Kolburan Geçer, A. (2021). Chatbot application in a 5th grade science course. Education and Information Technologies, 26(5), 6241–6265.

Torun, B., & Karamustafaoğlu, O. (2025). Fen öğretiminde ChatGPT kullanımı hakkında öğrenci görüşleri. Cumhuriyet Uluslararası Eğitim Dergisi, 14(1), 240–249.

Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis. Humanities and Social Sciences Communications, 12(1), 1–21.

Xu, H., Gan, W., Qi, Z., Wu, J., & Yu, P. S. (2024). Large language models for education: A survey. arXiv Preprint. https://arxiv.org/abs/2405.13001

Yıldız Durak, H. (2023). Conversational agent-based guidance: Examining the effect of chatbot usage frequency and satisfaction on visual design self-efficacy, engagement, satisfaction, and learner autonomy. Education and Information Technologies, 28(1), 471–488.

Yıldız Durak, H., Eğin, F., & Onan, A. (2025). A comparison of human-written versus AI-generated text in discussions at educational settings: Investigating features for ChatGPT, Gemini and BingAI. European Journal of Education, 60(1), e70014.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? International Journal of Educational Technology in Higher Education, 16(1), 39.

Zhao, S., Blaabjerg, F., & Wang, H. (2020). An overview of artificial intelligence applications for power electronics. IEEE Transactions on Power Electronics, 36(4), 4633–4658.

Zhou, X., Teng, D., & Al-Samarraie, H. (2024). The mediating role of generative AI self-regulation on students' critical thinking and problem-solving. Education Sciences, 14(9), 1015.