

Büyük Dil Modelleriyle Kişilik İfadesi

Personality Expression with Large Language Models

Kerem Varnalı, Sinan Sonlu, Uğur Güdükbay

Bilgisayar Mühendisliği Bölümü, Bilkent Üniversitesi, Ankara, Türkiye

kerem.varnali@ug.bilkent.edu.tr, sinan.sonlu@bilkent.edu.tr, gudukbay@cs.bilkent.edu.tr

Özetçe —Bu çalışma, Büyük Dil Modellerindeki kişilik ifadesini incelemektedir. Bu amaçla, istenen kişilik özelliklerini ifade eden istemlere yönelik basit ve karmaşık yaklaşımların performansını karşılaştırılmıştır. Ayrıca, belirli bir kişinin adını ve bilgilerini eklemenin sonuçları iyileştirip iyileştirmediğini de incelenmiştir. Genel olarak, daha ayrıntılı istemlerin daha doğru sonuçlar elde ettiğini ancak kişilik faktörleri arasında bir korelasyon ortaya çıkarabileceği gözlemlenmiştir. Ayrıntılı istemler ayrıca daha uzun girdiyle sonuçlanır ve bu da üretken modeli şaşırtabilmektedir. Ayrılmış kişilik faktörlerine odaklanmak, korelasyonlardaki azalma nedeniyle daha iyi sonuçlar vermektedir.

Anahtar Kelimeler—Büyük Dil Modeli, Kişilik İfadesi, Beş Faktör Kişilik Kuramı.

Abstract—This study examines personality expression in Large Language Models. To this end, we compare the performance of naive and complex approaches to prompting to express the desired personality traits. We also examine whether including a specific person's name and information improves the results. Generally, we observe that more detailed prompts achieve more accurate results but could introduce a correlation between the personality factors. Detailed prompts result in longer input, which could confuse the generative model. Focusing on separated personality factors yields better results due to decreased correlations.

Keywords—Large Language Model, Personality Expression, Five-Factor Model of Personality.

I. GİRİŞ

Kişilik özelliklerinin metin üzerinden ifadesi ve analizi; psikoloji, doğal dil ve veri bilimi gibi çeşitli disiplinleri içeren bir işlemdir. Kişilik, bireyi oluşturan benzersiz yönlerin kategorize etmektedir; teoriler, farklı kişilik boyutlarını ve bu boyular arası ilişkileri çeşitli biçimlerde temsil etmektedir. En yaygın kişilik teorilerinden biri Beş Faktör Kişilik Kuramıdır. Bu çalışma, Büyük Dil Modelleri (Large Language Model veya LLM) tarafından oluşturulan metinlerde algılanan kişilik faktörlerini analiz etmek için Büyük Beş Envanteri (Big Five Inventory veya BFI) [1] kişilik tanımlarını kullanmaktadır.

İnsan benzeri iletişim özellikleri gösteren yapılar oluşturmayı amaçlayan Yapay Zeka sistemlerine yönelik artan ilgi ve çalışmalar, bu oluşumların psikolojik açıdan incelenmesini de gerekli kılmaktadır. Bu çalışmanın temel odağı, üretken sistemlerin gerçekçiliğini etkileyen ana unsurlardan biri olan algılanan kişilik özellikleridir. Konuşkan araçlarda, kişilik ifadesi etkileşim kalitesini ve kullanıcı katılımını artırır [2],

[3]. Benzer bir yaklaşım, metin tabanlı yöntemlerde, LLM yanıtlarını pekiştirme amacıyla kullanılabilir [4]. Bu çalışma, LLM aracılığıyla kişilik ifadeli metin üretme üzerinedir. LLM yanıtlarını şekillendirmedeki önemli bir unsur, modele girdi olarak sunulan istem metnidir. Basit istemler, belirli bir noktaya kadar çıktıyı istenilen yönde şekillendirebilirken, belirgin kişilik ipuçları üretmek için ayrıntılı istemler gerekmektedir.

LLM yanıtlarındaki kişilik ifadesi, kullanıcı çalışmaları yardımıyla ölçülebilir; ancak, güncel araştırmalar, modelleri değerlendirmek için yine LLM'lerden yararlanmaktadır. Yaygın yaklaşımlardan biri, BFI gibi kişilik anketi sorularının LLM'e yöneltilerek alınan cevaplar üzerinden kişilik puanının hesaplanmasıdır. Bazı çalışmalar, kişilik anketlerinin LLM'lere tam olarak genelleştirilemediğini öne sürse de [5], bu yaklaşım LLM kişiliğini analiz etmek için hızlı bir öncül sunmaktadır. Bu kısıtlı değerlendirme üzerinden başarılı bulunan bir sistemin takip eden bir kullanıcı çalışmasıyla da değerlendirilmesi, sonuçların isabetliliğini doğrulama açısından uygun olacaktır.

Bu çalışmada, ünlü bir şahsın isim ve bilgilerini istem olarak kullanmak da dahil olacak biçimde, LLM ile kişilik ifadesi üzerine çeşitli yaklaşımlar denenmiştir. LLM'ler, şahısları yansıtmada başarılıdır [6]–[8], bu sebeple, temsil edilen şahsın kişilik yönünü ek istemler aracılığıyla pekiştirmek amaçlanmıştır. Düşük, nötr ve yüksek kişilik özelliklerini temsil etme yönünde çeşitli yaklaşımların başarısı karşılaştırılmıştır.

II. İLGİLİ ÇALIŞMALAR

A. Büyük Dil Modelleri

LLM yapıları kapsamlı eğitim kümeleri kullanarak otonom aracı işlemlerine elverişli çözümler sunar [9]. Uygulamalar, finans sektöründen [10] sağlığa [11] çeşitli konularda LLM yapısından yararlanmaktadır. Mevcut Yapay Zeka araştırmaları, farklı ortamlardan üretken modelleri birleştirmeyi amaçlamaktadır ve LLM yapısı bu tarz karmaşık denetleyici rolleri üstlenmeye uygundur [12]. Örneğin, LLM yapıları metin tabanlı kullanıcı girdisini yorumlayabilir ve ilgili alt sistemleri kullanıcının isteğine yanıt vermek üzere tetikleyebilir. Bu noktada, LLM'lerin isabetliliğini iyileştirmek çoklu ortam çıktıları hedefleyen sistemler için esastır.

B. Kişilik Özelliklerine Odaklanan Sistemler

İnsanlar, dijital içeriklerde farklı kişilikler gözlemleme eğilimindedir [13]. Bir robot veya yapay bir aracı, kişilik ifadesini amaçlamasa bile, insanlar belirli görünüm veya davranışlarla ilgili unsurları kişilik özellikleriyle ilişkilendirir; bu sebeple, belirli kişilik faktörlerinin ifadesini hedeflemek, dijital karakterlerin daha inandırıcı hale gelmesini sağlar [14]. Çalışmalar

fiziksel görünüm [15] veya el hareketleri [16] kullanarak dijital karakterlerde kişilik ifadesini amaçlamışlardır. Genellikle, farklı görsel-işitsel özellikleri birlikte kullanan çok ortamlı yaklaşımlar, farklı kişilik özelliklerini yansıtmaya ve algılama konusunda daha iyi sonuçlar vermektedir. İnsan iletişiminin farklı yönleri, kişiliği işaret eden ipuçlarını içerir; örneğin, bir konuşma sırasında ses ve tonlama, konuşulan kelimeler, insanların kişilik özelliklerini nasıl gözlemlediğini etkiler. Sistemler, çeşitli uygulamalarda farklı kişilik türlerini yansıtmaya amacıyla bu tür ipuçlarını tespit edebilir veya taklit edebilir.

C. Metinde Kişilik İfadesi

Metin tabanlı kişilik ifadesinde önemli bir unsur kelime seçimi ve dil bilgisi kullanımıdır [17]. Örneğin, kullanılan kelimelerin karmaşıklığı, dilin resmiyeti ve sözcüklerin duygusal içerikleri diyalogda farklı kişilik özelliklerini ifade etmek için kullanılabilir [18]. Önceki çalışmalar, elle hazırlanmış kelime ve cümle manipülasyonlarına dayanırken, güncel yaklaşımlar, sistemlere kişiliği tanıtmak için belirli istemleri veya etiketli veri kümelerini kullanmaktadır.

Basit yaklaşım, "Siz dışa dönük bir kişisiniz" veya "Dışa dönük gibi davranın" tarzı istem mesajları kullanarak, amaçlanan genel kişilik özelliklerini belirtmektir. Bu, özellikle eğitim kümesi kişilik tabanlı içeriklere dair yeterli miktarda örnek içermediğinde, karışık sonuçlara yol açacaktır. Daha iyi sonuçlar elde etmek için, amaçlanan kişilik faktörleri ve bu faktörlerin alt parçalarına ilişkin açıklayıcı kelimeler kullanılmalıdır. Kişilik istemi (Personality Prompting veya P^2) [19], farklı üst düzey kişilik özelliklerinin daha uzun ve açıklayıcı istemlere eşleştirildiği bir yöntemdir. Bu kişilik eşlemesi, farklı kişilik özelliklerini daha iyi ifade etmeyi sağlamaktadır.

III. YÖNTEM

A. Kişilik Analizi

LLM kişilik analizi, mevcut literatürdeki öz değerlendirme kişilik analizine benzer şekildedir. İnsan kişilik analizi, genellikle öz bildirimlere veya akran bildirimlerine dayalı her kişilik kategorisini ölçen anket soruları üzerinden gerçekleştirilir. Farklı kişilik anketleri, temel alınan kişilik teorisi ve soru sayısı bakımından çeşitlilik göstermektedir. Bu çalışmada, kişiliği daha ayrıntılı olarak ölçmek amacıyla 120 soru içeren BFI anketi kullanılmıştır. Zaman açısından kısıtlı çalışmalarda kullanılmak üzere daha kısa BFI formları da mevcuttur; kullanıcı çalışmaları, genellikle, kullanıcı yorgunluğunu önlemek amacıyla kısa formdaki anketleri kullanmaktadır. Ancak, LLM hızlı yanıtlar üretebildiğinden, uzun formlu anketlerin kullanımını zorluk teşkil etmez ve anketin isabetliliğini artırır. Mevcut kişilik anketleri, insan kişiliğini ölçmeye odaklıdır; öte yandan, gelecek çalışmalarda LLM kişiliğine odaklanan tarzda sorular oluşturularak daha verimli bir analiz sağlanabilir.

BFI, Beş Faktör Kişilik Kuramının boyutlarını tanımlayan sıfatlar içerir; bunlar Açıklık (Openness), Sorumluluk (Conscientiousness), Dışa Dönüklük (Extraversion), Uyumluluk (Agreeableness) ve Nevrotiklik (Neuroticism) olmak üzere beş ana faktör üzerinden belirlenmiştir. Faktörlerin ilk harfleri OCEAN kısaltmasını oluşturmaktadır, bu sebeple bu model OCEAN olarak da anılmaktadır. Söz konusu kişilik faktörlerinin adları tanımlayıcıdır, ancak her faktörü ifade eden alt özelliklerin bilinmesi, ilgili faktörlerin daha etkin biçimde temsili için önem

arz etmektedir. Listelenen sıfatlar [20], kişilik faktörlerini daha ayrıntılı tanımlamaya yardımcı olacaktır:

- **Açıklık:** Sanatsal, meraklı, hayal gücü geniş, anlayışlı, özgün, geniş ilgi alanları sahibi.
- **Sorumluluk:** Verimli, düzenli, planlı, güvenilir, sorumlu, titiz.
- **Dışa dönüklük:** Aktif, iddialı, enerjik, coşkulu, dışa dönük, konuşkan.
- **Uyumluluk:** Takdir edici, affedici, cömert, nazik, sempatik, güvenen.
- **Nevrotiklik:** Kaygılı, kendine acıyan, gergin, alıngan, dengesiz, endişeli.

Bahsi geçen sıfatlar, kişilik faktörlerini alt düzey özellik kategorileriyle eşleştiren P^2 yöntemi ile kullanılmaktadır. BFI ile LLM kişiliği ölçerken, Tablo I'deki yaklaşım kullanılmıştır. Bu noktada, LLM'in sorulara cevapları üzerinden ortalama skor değerleri (S) ve standart sapma (σ) hesaplanmaktadır.

TABLO I: LLM kişilik analizi test mekanizması [19].

Şablon
İstem: {Değişken 1}
Şu ifade üzerinden: {Değişken 2}, seni hangi seçenek daha iyi temsil eder:
Seçenekler: (A). Çok isabetli / (B). Az isabetli
(C). Ne isabetli ne isabetsiz / (D). Az isabetsiz / (E). Çok isabetsiz
İfade Örnekleri
Hayal kurmada güçlük çekerim ($-O$)
Davalar hakkında tutkuluymumdur ($+O$)
Sıklıkla son dakika planları yaparım ($-C$)
Benden beklenenden fazlasını yaparım ($+C$)
Olayları kendi akışına bırakırım ($-E$)
İnsanların yanında rahat hissederim ($+E$)
Çoğu sorunun cevabını kendim biliyor olurum ($-A$)
Başkalarına yardım etmekten hoşlanırım ($+A$)
Nadiren aşırılıklarım olur ($-N$)
Sonradan pişman olduğum şeyler yaparım ($+N$)

Tabloda, şablon olarak etiketlenen metin, mevcut istem üzerinden LLM'in kişilik özelliklerini tespit için girdi olarak kullanılmaktadır. Tabloda iki değişken bulunmaktadır; ilk değişken, LLM kişiliğini belirlemede rol oynayan istem girdisidir. Bu istem girdisi, kullanılan yöntemle göre basit veya detaylı olabilir. İkinci değişken ise, ölçülen her kişilik faktörüne göre değişen kişilik anketi sorularıdır. Bu sorular için Uluslararası Kişilik Ögesi Havuzu (International Personality Item Pool veya IPIP-NEO-120) kullanılmıştır [21]. Her soru bir kişilik faktörüyle doğru veya ters orantılıdır, orantının işareti ilgili ölçümün pozitif veya negatif oluşunu belirler.

Örneğin, soru, "İnsanların yanında rahat hissederim" gibi olumlu bir özellik ifade ediyorsa verilen cevabın işareti pozitifdir veya "Hayal kurmada güçlük çekerim" gibi olumsuz bir soruysa işaret negatiftir. LLM'in verdiği cevaba göre çok isabetsizden çok isabetliye doğru 1'den 5'e yükselen sayısal değerlere karşılık gelen cevaplar mevcuttur. Her soru için alınan cevap, soru ile ilişkilendirilen pozitif veya negatif işaret göz önünde bulundurularak, kişilik faktörleri için ortalama ve standart sapma değerleri hesaplanır. 120 soru için 30 farklı çalıştırma üzerinden, d ile temsil edilen farklı kişilik faktörleri için aşağıdaki denklem üzerinden (S) ve (σ) hesaplanır:

$$\bar{S}_d = \frac{1}{30} \sum_{j=1}^{30} (S_d)_j \quad \sigma = \sqrt{\frac{\sum_{i=1}^{30} ((S_d)_i - \bar{S}_d)^2}{30}}$$

Tablolarda belirtilen insan ortalamasını elde etmek için IPIP-NEO-120 verileri kullanılmıştır. Bu veri kümesinde, toplam 619,150 yanıt üzerinden katılımcılar için geçerli ortalama OCEAN skor ve standart sapma değerleri hesaplanmıştır.

B. Kişilik Ekleme Yöntemleri

Temel olarak iki farklı kişilik ekleme yöntemine odaklanılmıştır (bkz. Şekil 1). İlk yöntem, yalnızca ilgili kişilik faktörünün adının istem olarak kullanıldığı basit yaklaşımdır. İstem, ilgili faktör için amaçlanan değerin düşük, yüksek veya nötr oluşuna göre biçimlendirilir. Öte yandan, P^2 yaklaşımı, istemin daha ayrıntılı bir şekilde yansıtılmasıyla elde edilir. Bu yaklaşımda, her faktör için tanımlanan sıfatlar amaçlanan kişiliği daha ayrıntılı tarif etmektedir. Ek olarak, belirli bir şahsın isminin ve bilgilerinin kullanımının amaçlanan kişiliği pekiştirmesi öngörülmektedir. Şekilde belirtilen girdilerin farklı birleşimlerinin ölçülen kişiliğe etkisi gözlemlenmiştir.



Şekil 1: Çalışmada kullanılan istem çeşitlerinin özet akışı.

C. Tekil Kişilik Faktörlerini İfade

Öncelikle, ölçülen kişilik özellikleri üzerindeki etkisini tespit amacıyla tekil kişilik faktörlerine odaklanılarak istemler oluşturulmuştur. Hedef kişilik faktörü için istenilen değer yüksek olarak belirtilirken geri kalan faktörler nötr olarak belirtilmiştir, böylece yalnızca hedef faktörlerin yüksek olması amaçlanmıştır; elde edilen sonuçlar Tablo II'de gösterilmiştir.

TABLO II: Tekil yüksek faktörleri ifade basit yaklaşım.

Hedef	Açıklık		Sorumluluk		Dışa Dönüklük		Uyumluluk		Nevrotiklik	
	S	σ	S	σ	S	σ	S	σ	S	σ
O	4.08	1.22	3.58	1.47	3.82	1.21	4.04	0.89	3.41	1.70
C	3.79	1.22	4.08	1.04	3.33	1.06	3.96	0.89	3.21	1.55
E	3.96	1.24	3.58	1.29	3.83	1.46	3.92	1.03	3.42	1.63
A	3.86	1.20	3.92	1.19	3.70	1.37	4.04	1.02	3.42	1.63
N	3.80	1.2	4.08	1.04	3.25	1.20	3.88	0.97	4.08	1.59
Nötr	3.54	1.51	3.46	1.12	3.13	0.83	3.71	0.98	3.71	1.51
İnsan	3.44	1.06	3.60	0.99	3.41	1.03	3.66	1.02	2.80	1.03

Bu noktada amaçlanan, hedef kişilik değerini yüksek olarak ifade ederken geri kalan faktörlerin nötr oluşudur. Bu sebeple, her faktörün nötr ifadesi amaçlandığı zaman elde edilen değerler bir referans olarak kullanılmak üzere belirlenmiştir. Hedef faktör açısından değişikliklerin, geri kalan faktörlerin nötr oluşunu bozmaması beklenmektedir. Ancak tabloya göre bu durum her zaman beklenen şekilde değildir; Açıklık, Dışa Dönüklük ve Uyumluluk faktörlerinin birbirleriyle ilişkili olarak değişimi gözlemlenmektedir. Benzer şekilde, Sorumluluk ve Nevrotiklik birbirleriyle ilişkilidir. Ayrıca, Yüksek Uyumluluğun Sorumluluk üzerinde etkisi olduğu gözlemlenmektedir. Sıradaki deney P^2 yönteminin etkisini ölçmektedir.

P^2 yöntemi ile standart sapma insana benzer değerlere yaklaşmaktadır ve hatta bazı durumlarda daha az sapma göstermektedir. İlk deneye benzer şekilde her faktörü nötr olarak

TABLO III: Tekil yüksek faktörleri ifade P^2 yaklaşımı.

Hedef	Açıklık		Sorumluluk		Dışa Dönüklük		Uyumluluk		Nevrotiklik	
	S	σ	S	σ	S	σ	S	σ	S	σ
O	4.08	0.86	4.04	0.35	3.63	1.03	4.21	0.50	2.91	0.93
C	3.46	0.91	4.38	0.48	3.17	0.94	4.08	0.76	2.04	0.68
E	3.83	0.69	4.13	0.44	4.29	0.84	4.13	0.88	2.00	0.71
A	3.71	0.74	4.29	0.46	3.08	1.08	4.71	0.46	2.79	1.04
N	3.67	0.75	3.96	0.54	3.58	0.81	4.00	0.50	3.71	0.74
Nötr	3.67	0.69	4.04	0.35	3.13	0.97	4.08	0.57	2.21	0.76
İnsan	3.44	1.06	3.60	0.99	3.41	1.03	3.66	1.02	2.80	1.03

ifade eden bir durum referans olarak kullanılmıştır. Tablo III'de, ilk deneye benzer olarak Açıklık ve Dışa Dönüklük arasında korelasyon bulunmaktadır. Benzer şekilde Yüksek Uyumluluk, Yüksek Sorumluluk üzerinde etki sahibidir. Ancak, ilk deneyden farklı olarak, Nevrotiklik ve Sorumluluk korelasyonlu bulunmamaktadır ve Uyumluluk, Açıklık ve Dışa Dönüklük üzerinde etki sahibi değildir.

D. Belirli Şahsın Kişiliğinin Eklenmesi

Bu deneyde tanınmış bir kişinin özelliklerini ifade amacıyla basit ve P^2 yaklaşımları kullanılmıştır. Bu noktada, isabetliliği artırmak adına ek değişkenlere de yer verilmiştir. Tanınmış kişi hakkında isabetli sonuçlar verebilmesi için, LLM'in eğitimi sırasında bu kişi hakkında yeterli miktarda veri kullanılmış olmalıdır. Bu sebeple, internet üzerinde hakkında büyük miktarda bilgi bulunan Elon Musk tercih edilmiştir. Elon Musk'ın kişilik özelliklerine dair de bilgi mevcuttur: Yüksek Açıklık, Düşük Sorumluluk, Nötr Dışa Dönüklük, Yüksek Uyumluluk, ve Düşük Nevrotiklik. İstem olarak öncelikle Elon Musk'ın ismi ve hakkındaki bilgiler girdi verilerek bir deney yapılmıştır.

TABLO IV: İsim ve bilgi üzerinden şahsın kişilik ifadesi.

Küme	Açıklık		Sorumluluk		Dışa Dönüklük		Uyumluluk		Nevrotiklik	
	S	σ	S	σ	S	σ	S	σ	S	σ
İsim	4.13	1.05	4.00	1.22	4.50	0.71	3.79	1.32	2.13	1.05
Bilgi	4.04	1.51	4.13	1.33	4.34	1.28	4.13	1.51	2.79	1.71

Bu noktada, şahsın yalnızca ismi veya hakkındaki genel bilgiler kullanıldığı zaman, şahsın için belirlenen kişilik özelliklerine ulaşılmadığı gözlemlenmektedir (Tablo IV). Bu sebeple, isim ve bilgi girdilerine ek olarak hedef kişilik özelliklerinin de istemin bir parçası olarak sunulduğu bir yaklaşım denenmiştir. Bu noktada basit ve P^2 yöntemleri ayrıca işlenmiştir.

TABLO V: Şahıs girdisinin kişilik istemiyle birleştirilmesi.

Küme	Açıklık		Sorumluluk		Dışa Dönüklük		Uyumluluk		Nevrotiklik	
	S	σ	S	σ	S	σ	S	σ	S	σ
Basit	4.13	1.27	3.75	1.56	4.04	1.37	4.21	1.00	3.54	1.63
Basit + İsim	4.17	0.80	4.20	0.81	4.17	0.86	4.04	0.73	1.91	0.76
Basit + Bilgi	4.08	1.32	3.75	1.45	4.17	1.49	4.25	1.33	2.92	1.61
P^2	4.18	0.70	4.38	0.70	3.42	1.53	4.70	0.46	1.89	0.93
P^2 + İsim	4.33	0.85	4.26	0.50	3.25	1.40	4.58	0.49	2.38	1.22
P^2 + Bilgi	4.33	0.90	4.09	0.98	3.21	1.47	4.83	0.37	1.79	0.82

Tablo V'de gözlemlendiği üzere, basit yaklaşım tek başına Yüksek Nevrotiklik göstermektedir ancak isim eklenmesiyle Nevrotiklik değeri olması gerektiği şekilde düşük hale gelmektedir. Öte yandan, basit yaklaşım ile isim veya bilgi kullanılması Sorumluluk ve Dışa Dönüklük açısından istenilen değerleri vermemektedir. Bu sebeple, aynı yaklaşım P^2 yöntemiyle de kullanılmıştır.

P^2 yaklaşımı ile daha isabetli sonuçlar elde edilmiştir. P^2 yöntemi sapmanın düşürülmesinde yardımcı olmuştur ve kişilik faktörleri şahsın değerlerine daha yakın hale gelmiştir. Şahıs hakkında bilgi eklenmesiyle sonuçlar, basit yöntemde elde edilene kıyasla, daha yüksek oranda iyileşme göstermektedir. Öte yandan, Sorumluluk açısından sonuçlar görece kötüdür.

Bu noktaya kadar her faktör için girdiler birleştirilerek verilmiştir. Sonuçları daha iyi karşılaştırma amacıyla bu istemlerin ayrı olarak verildiği durumların sonuçları da karşılaştırılmıştır. Bu noktada, Elon Musk'ın kişilik özelliklerini ayrı ve birleşik olarak verilmesiyle elde edilen değerler incelenmiştir.

TABLO VI: Ayrık ve birleşik girdi performansları.

Küme	Açıklık		Sorumluluk		Dışa Dönüklük		Uyumluluk		Nevrotiklik	
	S	σ	S	σ	S	σ	S	σ	S	σ
Basit Ayrık	4.10	1.10	2.12	1.21	3.88	1.05	4.46	0.87	2.12	0.82
Basit Birleşik	4.13	1.27	3.75	1.56	4.04	1.37	4.21	1.00	3.54	1.63
P^2 Ayrık	4.54	0.76	1.46	0.50	3.29	0.93	5.00	0.00	1.46	0.50
P^2 Birleşik	4.18	0.70	4.38	0.70	3.42	1.53	4.70	0.46	1.89	0.93

Tablo VI'e göre, kişilik faktörlerinin sisteme birleştirilerek sunulması, faktörler arası korelasyona sebep olup sonuçları istenilen değerlerden biraz uzaklaştırmaktadır. Öte yandan, faktörlerin ayrı olarak kullanımı sonuçları iyileştirmektedir.

IV. SONUÇ

Bu bildiri, LLM ile kişilik ifadesi konusunda kullanılabilecek yöntemleri karşılaştırmaktadır. Kişilik kazandırmada basit ve P^2 yaklaşımları test edilmiştir. P^2 yöntemi, daha ayrıntılı bir istem desteklemesi sebebiyle daha iyi sonuç vermektedir. Ancak, tek bir istemle beş faktöre dair hedef değerleri kazandırmaya çalışmak isabetliliği düşürmektedir. P^2 yönteminde birden fazla istem verirken, sıfatlardan bazılarının diğerleri üzerinde etkisi olması ve tek bir faktör açısından iyi sonuçlar alınsa bile, diğer faktörlerin istenilen değerlerden uzaklaşmasına sebep olmaktadır. Öte yandan, basit yaklaşım, açıklamaların istenilen faktörleri yeterli tarif edememesi sebebiyle başarısız bulunmuştur. Bir başka sorun, kişilik eklerken tüm P^2 tanımlarının birleştirilmesidir; bu şekilde LLM'e sunulan istem, tek bir faktöre odaklanılan duruma kıyasla 5 kat uzundur. Bu durum, uzun istem dolayısıyla üretilen çıktının tutarsızlık göstermesine yol açabilmektedir.

Bu noktada, hedeflenen kişilik özelliklerinin belirli bir şahsın ismi ve bilgileriyle birlikte kullanımının, hedeflenen kişilik değerlerinin ifadesi konusunda sonuçları iyileştirdiği gözlemlenmiştir. Amaçlanan kişilik değerlerini LLM kullanarak yansıtmada konusunda daha isabetli yöntemler araştırmamız sürmektedir. Bu konuda, kullanıcı girdisini şekillendirmede kullanılan farklı istem modelleri umut vadetmektedir. Çalışma kapsamında tek şahıs üzerine odaklanılmış olup, deneylerin farklı şahıslar için tekrarlanarak yöntemin genel etkisinin ölçülmesi planlanmaktadır. LLM'in eğitim sürecinde kişilik etiketli veri kullanılarak da sonuçların pekiştirilmesi mümkündür. Bu bildirinin bilgisayar tabanlı kişilik ifadesi konusunda gelecek çalışmalara ilham olması umulmaktadır.

KAYNAKLAR

- [1] O. P. John, E. M. Donahue, and R. L. Kentle, *The Big Five Inventory versions 4a and 5*. University of California, Berkeley, Institute of Personality and Social Research, 1991.
- [2] G. Ball and J. Breese, "Emotion and personality in a conversational agent," in *Embodied Conversational Agents*. Cambridge, MA, USA: MIT Press, 2001, p. 189–219.
- [3] S. T. Völkel, R. Schödel, D. Buschek, C. Stachl, V. Winterhalter, M. Bühner, and H. Hussmann, "Developing a personality model for speech-based conversational agents using the psycholexical approach," in *Proceedings of the CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–14.
- [4] K. Pan and Y. Zeng, "Do LLMs possess a personality? making the MBTI test an amazing evaluation for large language models," *arXiv preprint arXiv:2307.16180*, 2023.
- [5] F. Dörner, T. Sühr, S. Samadi, and A. Kelava, "Do personality tests generalize to large language models?" in *Socially Responsible Language Modelling Research*, 2023.
- [6] T. Hu and N. Collier, "Quantifying the persona effect in LLM simulations," *arXiv:2402.10811 [cs.CL]*, 2024.
- [7] G. Sun, X. Zhan, and J. Such, "Building better AI agents: A provocation on the utilisation of persona in LLM-based conversational agents," in *Proceedings of the 6th ACM Conference on Conversational User Interfaces*, 2024, pp. 1–6.
- [8] Y. Choi, E. J. Kang, S. Choi, M. K. Lee, and J. Kim, "Proxona: Supporting creators' sensemaking and ideation with LLM-Powered audience personas," *arXiv:2408.10937v3 [cs.HC]*, 2025.
- [9] L. Wang, C. Ma, X. Feng, Z. Zhang, H. Yang, J. Zhang, Z. Chen, J. Tang, X. Chen, Y. Lin *et al.*, "A survey on large language model based autonomous agents," *Frontiers of Computer Science*, vol. 18, no. 6, p. 186345, 2024.
- [10] S. Wu, O. Irsoy, S. Lu, V. Dabrovolski, M. Dredze, S. Gehrmann, P. Kambadur, D. Rosenberg, and G. Mann, "BloombergGPT: A large language model for finance," *arXiv:2303.17564v3 [cs.LG]*, 2023.
- [11] X. Yang, A. Chen, N. PourNejatian, H. C. Shin, K. E. Smith, C. Parisien, C. Compas, C. Martin, A. B. Costa, M. G. Flores *et al.*, "A large language model for electronic health records," *NPJ digital medicine*, vol. 5, no. 1, p. 194, 2022.
- [12] S. Song, X. Li, S. Li, S. Zhao, J. Yu, J. Ma, X. Mao, and W. Zhang, "How to bridge the gap between modalities: Survey on multimodal large language model," *arXiv:2311.07594 [cs.CL]*, 2025.
- [13] Y. Moon and C. Nass, "How 'real' are computer personalities? psychological responses to personality types in human-computer interaction," *Communication Research*, vol. 23, no. 6, pp. 651–674, 1996.
- [14] F. de Rosi and C. Castelfranchi, "How can personality factors contribute to make agents more believable," in *Proceedings of the 13 Spring Days Workshop on Behaviour Planning for Lifelike Characters and Avatars*, vol. 8080, 1999.
- [15] L. P. Naumann, S. Vazire, P. J. Rentfrow, and S. D. Gosling, "Personality judgments based on physical appearance," *Personality and Social Psychology Bulletin*, vol. 35, no. 12, pp. 1661–1671, 2009.
- [16] M. Neff, Y. Wang, R. Abbott, and M. Walker, "Evaluating the effect of gesture and language on personality perception in conversational agents," in *Proceedings of the 10th International Conference on Intelligent Virtual Agents*, ser. IVA'10. Berlin, Heidelberg: Springer-Verlag, 2010, p. 222–235.
- [17] F. Mairesse, M. A. Walker, M. R. Mehl, and R. K. Moore, "Using linguistic cues for the automatic recognition of personality in conversation and text," *Journal of Artificial Intelligence Research*, vol. 30, pp. 457–500, 2007.
- [18] F. Mairesse and M. Walker, "PERSONAGE: Personality generation for dialogue," in *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*. Association for Computational Linguistics, 2007, pp. 496–503.
- [19] G. Jiang, M. Xu, S.-C. Zhu, W. Han, C. Zhang, and Y. Zhu, "Evaluating and inducing personality in pre-trained language models," *arXiv preprint arXiv:2206.07550 [cs.CL]*, 2023.
- [20] R. R. McCrae and O. P. John, "An introduction to the five-factor model and its applications," *Journal of Personality*, vol. 60, no. 2, pp. 175–215, 1992.
- [21] J. A. Johnson, "Measuring thirty facets of the five factor model with a 120-item public domain inventory: Development of the IPIP-NEO-120," *Journal of Research in Personality*, vol. 51, pp. 78–89, 2014.