

# TEKNOFEST

## HAVACILIK, UZAY VE TEKNOLOJİ FESTİVALİ

### EĞİTİM TEKNOLOJİLERİ YARIŞMASI

#### PROJE DETAY RAPORU

**PROJE ADI:** Eğitim Amaçlı Kural Tabanlı Türkçe Doğal dil işleme ve İnsansı robotta uygulaması

**TAKIM ADI:** İSTE Doğal Dil İşleme

**TAKIM ID:** T3-13447160

**TAKIM SEVİYESİ:** Üniversite-Mezun

**DANIŞMAN ADI:** -

## İçindekiler

### 1. Proje Özeti (Proje Tanımı)

Projede, Türkçe doğal dil işleme teknikleri ile tamamen Türkçe dilbilgisi kurallarına uygun, etkileşimli soru cevap diyalog uygulaması oluşturulacaktır. Geliştirilen uygulama, Otizmliler çocukların iletişim becerilerinin geliştirilmesi ve yabancı uyruklu öğrencilere Türkçe'nin etkileşimli bir şekilde öğretilmesi için insansı robot (NAO robot) üzerinde gerçekleştirilecek ve uygulamaya koyulacaktır. Özgün bir veri seti kullanılacaktır. Kullanılacak olan veri setinde tamamen soru-cevap cümlelerine odaklanılacak ve hedef kitleye özgün senaryolar ile desteklenecektir. Bu kapsamda hazırlanacak olan veri seti üzerinde normalizasyon işlemleri uygulanacak ve soru cevap ilişkilendirme yapılacaktır. Bu sayede veri setinde olmayan sorulara bile ilişkilendirme yapılarak cevap verilebilecektir. Uygulama alanı olarak kullanılacak Nao robot için özgün bir uygulama arayüzü geliştirilecek ve uygulama sepeti mantığıyla arayüz tüm öğretmenlerin kullanımına hazır hale getirilecektir. Ayrıca insansı robotun olmaması durumu için de animasyonlu simülasyon çalışması da yapılacaktır.

### 2. Problem/Sorun:

Otizmliler bireylerin, özellikle de otizmliler çocukların eğitiminde kullanılan sınırlı sayıda materyal vardır. İletişim becerilerinin gelişmesine katkı sağlayacak bu materyallerin başarımı otizmliler bireylerin ilgilerini çekmekle doğru orantılıdır. Geliştirilecek Diyalog uygulaması ve insansı robotta gerçekleştirilmesi ile bu bireylerin ilgisi yüksek derecede artacak olup eğitimlerine katkı sağlayacaktır. Ayrıca yabancı uyruklu öğrencilere (özellikle çekingen, utangaç, yeni gelmiş vb. insan etkileşimi kuramayan/kurmayan Erasmus öğrencileri ve Suriye'li öğrenciler vb. gibi) Türkçe'nin daha kolay ve interaktif bir şekilde öğretilmesi gerekmektedir.

### 3. Çözüm

Robotun özellikle otizmliler çocuklar için eğitim amaçlı kullanılması öngörülmüştür. Bu proje çalışması sonucunda daha önce birçok alanda uygulamalarda çalışılan Nao insansı robotun (Şekil 1) hareketlerinin kontrol edilmesi ve geliştirilmesi (uygulama kolaylığı açısından özgün uygulama arayüzü oluşturulması) ve ayrıca Türkçe doğal dil işleme tekniklerinin eklenmesi ile birlikte oluşturulacak Türkçe diyalog uygulaması ile otistik çocukların ilgisini çekebilecek, iletişim becerilerini arttıracak bir uygulama hayata geçirilmiş olacaktır. Ayrıca yabancı uyruklu öğrencilere (özellikle Erasmus öğrencileri ve Suriye'li öğrenciler vb.) Türkçe'nin daha kolay ve interaktif bir şekilde öğretilmesi sağlanacaktır.



Şekil 1. Nao İnsansı Robot

Sorun	Çözüm	Eğitimdeki Katkısı
<p>Otizmlili bireylerin iletişim becerilerinin geliştirilmesinde eğitim materyali eksikliği bulunmaktadır. Ayrıca yabancı uyruklu öğrencilerin özellikle pandemi sürecinde de görüldüğü gibi insan ilişkisi olmadan interaktif bir şekilde dil öğrenmeleri için eğitim materyali eksikliği bulunmaktadır.</p>	<p>Türkçe Doğal dil işleme teknikleri ile İnsansı robot kullanılarak Otizmlili bireylerin iletişim becerilerinin yükseltilmesi ve yabancı uyruklu öğrencilerin insan iletişimi olmadan robot ile interaktif şekilde dil öğrenimi sağlanacaktır.</p>	<p>Otizmlili bireylerin iletişim becerilerinin yükseltilmesi için eğitim materyali oluşturulmuş olacaktır ve yabancı uyruklu öğrencilerin insan iletişimi olmadan robot ile interaktif şekilde dil öğrenimi için bir uygulama alanı sağlanmış olacaktır.</p>

#### 4. Yöntem

Türkçe' de eş anlamlı veya yan anlamlı birçok kelime yer aldığından, elde edilen veri kümelerinin her birisi için ayrı ayrı sözcük analizi yapılması gerekmektedir. Eş anlamlı kelimelerin cümledeki anlamına göre değerlendirilmesi için mevcut algoritmalarından yararlanılması düşünülmektedir. Bu algoritmaların yeterli görülmediği noktalarda yeni ve özgün yöntemler geliştirilmesi planlanmaktadır. Sözcüklerin kök halinde veya ek almış halde olduğunu analiz edip türüne göre sınıflandırarak en uygun hale getirilmesi hedefler arasındadır.

Genel olarak Doğal dil işleme teknikleri temel olarak 4 aşamada gerçekleşmektedir. Bu aşamalar; Algılama, yöneltme, işleme ve tepki verme olarak sınıflandırılabilir. Algılama, ses veya metin olarak alınan verileri bilgisayarın anlayabileceği şekilde gerekli düzeltmelerin yapılmasıyla gerçekleşir. Yöneltme ile gelen dış etkenler (ses, metin) veri olarak metine dönüştürülür ve bu metin anlamlı küçük birimlere dönüştürme yöntemi ile kelime öbeklerine ayrılır. Aynı zamanda çevrilen metinsel cümlede geçen noktalama işaretleri ve gereksiz boşluklardan temizlenir ve saf cümle analiz edilmek üzere hazır hale getirilir. Alınan veriler analiz edilir. Cümle öğelerine ayrılır ve önceden oluşturulmuş olan veri setlerinden yararlanır. Karşılaştırmalı analiz için Türkçe dilinin yapısal özelliklerine göre uyarlayacağımız “geliştirilmiş Kelime vektör dönüşüm” yöntemi kullanılarak, boyutsal bir analiz yönetimi geliştirilerek amacına uygun ve hedeflenir bir analiz yapmak mümkün olabilecektir. Yapılan analiz sonucunda bir çıktı oluşturulacaktır, bu da tepki verme yöntemi olarak nitelendirilebilir. Projede uygulama alanı olarak seçtiğimiz robotumuzun (Nao İnsansı Robot) karşı kullanıcıya tepki vermesi ile arada bir iletişimsel ağ kurulacaktır. Nao insansı robotta yaptığımız önceki çalışmalarımızda da geliştirdiğimiz uygulama arayüzüne ve öğretilmiş hareketlerine ek olarak kural tabanlı diyalog sistemi oluşturulacaktır. Özetle Bu projede öncelikle özgün ve geliştirilebilir Türkçe soru-cevap diyalog tabanlı veri seti oluşturulacaktır. Daha sonra Cümlelerden kelimelere ve köklere ayırma yöntemleri kullanılacaktır ve kelimelerin doğru ilişkilendirilebilmesi için normalizasyon çalışmaları yapılacaktır. Daha sonra ise ilişkilendirilmiş soru-cevaplar üzerinden kural tabanlı Türkçe soru-cevap diyalog uygulaması geliştirilecektir. Uygulama alanı olarak kullanılacak Nao robot için özgün bir uygulama arayüzü geliştirilecek ve uygulama sepeti mantığıyla arayüz, tüm eğitmenlerin kullanımına hazır hale getirilecektir.

##### 5. Yenilikçi (İnovatif) Yönü

İnsansı Robot üzerinde eğitim amaçlı tamamen Türkçe soru-cevap Diyalog uygulaması bulunmamaktadır. Bu açıdan özgün olan bu çalışma özgün Türkçe soru-cevap veri seti ile de literatüre katkı sağlayacaktır.

##### 6. Uygulanabilirlik

Proje başarıyla gerçekleştirildiği takdirde otizmli bireylerin (özellikle otizmli çocukların) iletişim becerilerine yüksek oranda katkı sağlayacak eğitsel bir uygulama ortaya çıkacaktır. Ayrıca ülkemize gelen yabancı uyruklu öğrenciler için Türkçe'nin daha kolay ve efektif bir şekilde öğretilmesi sağlanacaktır. Proje, hedef kitle üzerinde yapılacak testlerin sonuçlarına göre daha birçok farklı alanda (özellikle eğitim alanında) ve sosyal sorumluluk projelerinde kullanılabilir.

##### 7. Tahmini Maliyet ve Proje Zaman Planlaması

Projenin yazılım maliyeti bulunmamaktadır. Hali hazırda üniversitemizde bulunan Nao insansı robot kullanılacaktır. Proje farklı kurumlarda eğitim amaçlı olarak kulla-

nılmaya başlarsa Nao insansı robotun güncel piyasa değeri yaklaşık 92 bin TL civarındadır. Projemiz simülasyonla Robot olmadan da uygulanabilir ancak özellikle otizmlili bireylerin odaklanma problemleri için robotun olması önemlidir.

İP No	İş Paketlerinin Adı ve Hedefleri	Kim(ler) Tarafından Gerçekleştirileceği	Zaman Aralığı (...-.. Ay)	Başarı Ölçütü ve Projenin Başarısına Katkısı
1	Özgün Türkçe Soru-Cevap Veri Seti Hazırlanması	Burcu Katarancı, Ayda Akçay	0-1 Ay	Projenin temelini oluşturmaktadır. Veri setinin büyüklüğü ve kalitesi direkt olarak başarımlı oranlarına etki edecektir.
2	Kural tabanlı Türkçe soru-cevap diyalog uygulamasının hazırlanması	Kadir Tohma, Ulaş Umut Demirtaş, Tolga Can Aydemir, Bahadır Ökkeş Tıraş	1-3 Ay	Tamamen Türkçe olarak hazırlanacaktır.
3	Nao İnsansı Robotun kullanılması ve projeye entegre edilmesi	Kadir Tohma	2-4 Ay	Otizmlili bireylerde (özellikle otizmlili çocuklar) iletişim becerilerinin gelişimi ve yabancı uyruklu öğrencilerde Türkçe'nin öğretilmesi açısından oldukça önemlidir. Çünkü dikkat çekicidir ve odaklanma açısından önemlidir.
4	Simülasyon için animasyon uygulaması hazırlanması (Nao robot karakterine benzer nitelikte)	Kadir Tohma	4-5 Ay	Mevcuttaki Nao insansı robotun kullanılamaz duruma gelmesi durumunda b planı olarak kullanılacaktır.
5	Hedef Kitle üzerinde testler ve başarımlı oranları tespiti	Tüm takım üyeleri	5-6 Ay	Projenin hedef kitle üzerindeki başarımlı ve sonuçları bu iş paketinde belirlenecektir.

## 8. Proje Fikrinin Hedef Kitle (Kullanıcılar):

Projenin hedef kitle öncelikle otizmlili bireyler ve yabancı uyruklu öğrencilerdir. Otizmlili bireylerin iletişim becerilerinin yükseltilmesi ve yabancı uyruklu öğrencilerin Türkçe'yi daha kolay öğrenebilmesi hedeflenmiştir.

## 9. Riskler

Projemiz için en büyük risk kullanılan insansı robotun çalışamaz/kullanılamaz duruma gelmesidir. Ancak böyle bir senaryo ile karşılaşılırsa b planı olarak simülasyon üzerinde devam edilebilir.

## 10. Proje Ekibi

**Takım lideri:** Kadir Tohma

Adı Soyadı	Projedeki Görevi	Okul	Projeyle veya problemle ilgili tecrübesi
Kadir Tohma	Takım Lideri/ Yazılımsal geliştirmeler ve insansı robotla ilgili çalışmalar	İskenderun Teknik Üniversitesi	Doğal dil işleme ve insansı robotlarla ilgili çalışmaları/yayınları mevcuttur.
Ökkeş Bahadır Tıraş	Yazılımsal geliştirmeler ve veritabanı ilişkilendirmeleri	İskenderun Teknik Üniversitesi	Doğal dil işleme üzerine çalışmalar yapmaktadır.
Tolga Can Aydemir	Yazılımsal geliştirmeler ve veritabanı ilişkilendirmeleri	İskenderun Teknik Üniversitesi	Doğal dil işleme üzerine çalışmalar yapmaktadır.
Ulaş Umut Demirtaş	Yazılımsal geliştirmeler ve veritabanı ilişkilendirmeleri	İskenderun Teknik Üniversitesi	Doğal dil işleme üzerine çalışmalar yapmaktadır.
Ayda Akçay	Veri seti hazırlanması	İskenderun Teknik Üniversitesi	Doğal dil işleme üzerine çalışmalar yapmaktadır.
Burcu Katrancı	Veri seti hazırlanması	İskenderun Teknik Üniversitesi	Doğal dil işleme üzerine çalışmalar yapmaktadır.

## 11. Kaynaklar

Adalı, E. (2012). Doğal Dil İşleme. Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi, 5(2).

Aşliyan, R., Günel, K., & Yakhno, T. (2007, December). Detecting misspelled words in Turkish text using syllable n-gram frequencies. In International Conference on Pattern Recognition and Machine Intelligence (pp. 553-559). Springer, Berlin, Heidelberg.

Berk, G., Erden, B., & Güngör, T. (2018). Turkish verbal multiword expressions corpus. In 2018 26th Signal Processing and Communications Applications Conference (SIU) (pp. 1-4). IEEE.

Çakıroğlu, Ü., & Özyurt, Ö. (2006). Türkçe Metinlerdeki Yazım Yanlışlarına Yönelik Otomatik Düzeltme Modeli. Elektrik-Elektronik-Bilgisayar Mühendisliği Sempozyumu ve Fuarı (Eleco 2006), BURSA, TÜRKİYE, ss.7-9

Çetin, F. S., & Eryiğit, G. (2018). Türkçe Hedef Tabanlı Duygu Analizi İçin Alt Görevlerin İncelenmesi–Hedef Terim, Hedef Kategori Ve Duygu Sınıfı Belirleme. Bilişim Teknolojileri Dergisi, 11(1), 43-56.

Coşkun, N. (2013). Türkçe Tümcelerin Ögelerinin Bulunması (Doctoral dissertation, Fen Bilimleri Enstitüsü).

Dönmez, İ., & Adalı, E. (2015). Türkçe Tümce Çözümlemede Vektör Yaklaşımı. Afyon Kocatepe Üniversitesi Fen Ve Mühendislik Bilimleri Dergisi, 15(3), 1-11.

Ergün, K., Kubat, C., Çağıl, G., & Cesur, R. (2000). İnternet ortamındaki tüketici yorumlarından özet bilgi çıkarımı. *Sakarya University Journal of Science*, 17(1), 33-40.

Güneş, A., & Tantuğ, A. C. (2018). Turkish named entity recognition with deep learning. In 2018 26th Signal Processing and Communications Applications Conference (SIU) (pp. 1-4). IEEE.

Güzey, C. ve Oflazer, K., 1994. Spelling Correction in Agglutinative Languages, Bilkent University Department Of Computer Engineering and Information Systems Technical Report, BU-CEIS-94-01, Ankara, Turkey.

Kılıçaslan, Y., Özlem, U. Ç. A. R., Güner, E. S., & Kemal, B. A. L. (2006). Otistik Ve Zihinsel Engelli Çocuklar İçin Doğal Dil İşleme Tabanlı Bir Yardım Aracı: Bir Başlangıç Çalışması. *Trakya Üniversitesi Fen Bilimleri Dergisi*, 7(2), 101-108.

Küçük, D., & Arıcı, N. (2016). Türkçe için Wikipedia Tabanlı Varlık İsmi Tanıma Sistemi. *Politeknik Dergisi*, 19(3), 325-332.

Kutlugün, M. A., & Şirin, Y. (2018). Turkish meaningful text generation with class based n-gram model. In 2018 26th Signal Processing and Communications Applications Conference (SIU)(pp. 1-4). IEEE.

Oflazer, K. ve Solak, A., 1992. Parsing Agglutinative Word Structures And Its Application to Spelling Checking for Turkish, In Proceedings of the 15th International Conference On Computational Linguistics, Nantes, France, August 23-28, p. 39-45.

Oflazer, K., 1993. Two-level Description Of Turkish Morphology, In Proceedings of the Sixth Conference Of The European Chapter Of The Association For Computational Linguistics, Utrecht, Netherlands, April 1993.

Solak, A., 1991. Design And Implementation of A Spelling Checker For Turkish, M.S. Thesis, Bilkent University, Ankara.

Topçu, S., Şen, C., Amasyalı, M.F. (2012). Türkçe Sohbet Robotu. Akıllı Sistemler ve Uygulamaları, ASYU.

