

# Almayla Artırılmış Üretim Modeli ile Türkçe İş İlanı Üretimi

## Turkish Job Posting Generation with Retrieval-Augmented Generation Model

*Yazarlar Gizlenmiştir*

**Öz**—Teknolojinin hızla gelişmesi ile özellikle yazılım alanındaki hızlı gelişmelerden dolayı gelişmeleri takip etmekte zorlanan insan kaynakları ve alan uzmanları arasındaki iletişim boşluğu, eksik veya hatalı iş ilanlarının da yazılmasına sebep olmuştur. Yazılım iş ilanlarının daha hızlı ve doğru yayınlanması, ilgi çekici iş ilanı metni oluşturulması ve bu alanda Türkçe dilinde yapılan literatür çalışmasının olmaması bu çalışma için motivasyon kaynağı olmuştur. Hibrid Almayla Artırılmış Üretim (Retrieval-Augmented Generation - RAG), soru çoğaltma ve yeniden sıralama(rerank) sonucu elde edilen metinlerin faydalı sonuçlar ürettiği ortaya çıkmıştır. Uygulanan RAG sistemi ile kullanılan büyük dil modelinden %0.012 oranında daha gerçeğe yakın ve anlamlı metinler üretilebildiği gözlemlenmiştir. Çevrimiçi harici kaynaklardan elde edilen Türkçe yazılım iş ilanları ile ileri düzeyde doğal dil işleme teknikleri ve açık kaynak kodlu büyük dil modelleri kullanılarak literatüre bir katkı sağlanmak hedeflenmiştir.

**Anahtar Kelimeler**—Doğal Dil İşleme, RAG, Hibrid-RAG, İş İlanı, Büyük Dil Modelleri, Almayla Artırılmış Üretim

**Abstract**—With the rapid development of technology, the communication gap between human resources and field experts who have difficulty in following the developments, especially in the software field, has caused incomplete or incorrect job postings to be written. The motivation for this study was the faster and more accurate publication of software job postings, the creation of attractive job posts and the lack of literature studies in the Turkish language in this field. It was revealed that the texts obtained as a result of Hybrid Retrieval augmented generation (RAG), question multiplication and reranking produced useful results. It was observed that with the applied RAG system, texts that were 0.012% more realistic and meaningful could be produced than the large language model used. It was aimed to contribute to the literature by using advanced natural language processing techniques and large language models with Turkish software job postings obtained from external online sources.

**Keywords**—Natural Language Processing, RAG, Hybrid-RAG, Job Posting, Large Language Models, Retrieval-Augmented Generation

### I. GİRİŞ

Günümüzde internetin yaygın olarak kullanılmasıyla birlikte çevrimiçi iş ilanı yayınlama süreci çeşitli bilgi ve becerilere sahip iş arayan geniş kitlelere ulaşabilmesi açısından büyük bir avantaj sağlar. Ayrıca daha düşük maliyetli olması sebebi ile yaygın olarak tercih edilen bir yöntem olarak da öne çıkmaktadır. Dijital dönüşümün hızlanması ile şirketler dönüşüme ayak uydurmak ve rekabetçi ortamda varlığını sürdürmek için yenilikçiliğini artırmaya ve sürdürmeye sürekli olarak mecburdurlar. Hem

uluslararası hem de ülkemizde kullanılan çeşitli çevrimiçi iş ilanı yayınlama web siteleri bulunmaktadır. Adaylar bu iş ilanlarına herhangi bir zamanda ve yerde başvurular yapılabilir, aynı zamanda insan kaynakları da bu iş ilanlarını aynı şekilde çevrimiçi olarak ilan edilebilmektedir. Teknolojinin ve bilgisayar bilimlerinin ilerlemesinden dolayı bu alanlarda birçok iş ilanı yayınlanmaktadır. Bilgi teknolojileri alanındaki iş ilanlarının çevrimiçi olarak yayınlanması hem isabetli hem de tutarlı bir yaklaşımdır. Son yıllarda meydana gelen teknolojik gelişmelere bağlı olarak bilgi teknolojileri alanında değişen ihtiyaçların olması ile çeşitli ve yeni iş imkanları oluşmuştur. Bu gelişmelerden dolayı oluşan iş pozisyonlarında veya endüstrinin de gelişmesi ile birlikte halihazırda zaten var olan pozisyonlarda oluşan yüksek istihdam ihtiyacı, hızlı ve verimli olması sebebi ile birçok iş ilanının çevrimiçi yayınlanmasına zemin hazırlamıştır. Çevrimiçi iş ilanı yayınlamak, kolay ve hızlı başvuru imkânı tanınmasından dolayı ilan ile eşleşen adayın daha yüksek ihtimalle bulunması işverene büyük önemde katkı sağlamaktadır.

Doğal dil işlemenin gelişmesi ile birlikte geçmiş verilere dayanılarak metin üretimi, soru-cevap gibi çeşitli uygulamalar yaygınlaşmıştır. Günümüzde bu uygulamaları oluşturmak için RAG modelleri, ince ayarlama teknikleri, aktarım öğrenmesi gibi algoritmalar kullanılmaktadır. Doğal Dil işlemede kullanılan bu algoritmaların bazı ana kullanım alanları şunlardır: insan dilinin anlaşılması, makineler tarafından doğal dilde metin üretimi, çeviri ve soru-cevap sistemleridir. Üretken yapay zeka alanında yer alan doğal dil işlemede metin üretimi, 2017 yılında 'Attention Is All You Need'[1] makalesi ile tanıtılan dönüştürücülerin bağlamı anlama ve etrafındaki kelimeleri dikkate alma özellikleri sayesinde metin üretme görevleri etkili bir şekilde yapılır.

Dijital dönüşüm ile dijital teknolojiler kullanılmakta ve bu teknolojiler arasında doğal dil işleme de bulunmaktadır. Doğal dil işlemede yer alan ve üretken yapay zekanın bir alt dalı olan metin üretimi sayesinde birçok insan gücü gerektiren görevler otomatize edilmiş veya bu uygulamalar yapılan görevlere ve işlere yardımcı asistan olarak kullanılmıştır. Hızla gelişen ve büyüyen teknoloji alanında yazılım sektörü için iş ilanı yazmak insan kaynakları için zorlayıcı olabilmektedir. Bu tür iş ilanlarında alan özelinde kelime kullanımı fazla olduğu için veya kullanılan cümlelerin alan ile ilgili olmasından dolayı ve insan kaynaklarının da bu konularda doğal olarak yetersiz kalmalarından dolayı, bazen hatalı veya gereksiz anahtar kelime kullanımı veya cümle kullanımları ile karşılaşmaktadır. İnsan benzeri metin üretmek ve iş tanımı için en uygun iş ilanları hazırlamak önemli hale gelmiştir. Bu çalışmada tecrübesiz insan

kaynaklarına faydalı olabileceği gibi, tecrübeli insan kaynakları için de yol gösterici olabileceği düşünülmüştür. Literatürde Türkçe iş ilanı metni üretmek ile ilgili herhangi bir araştırmaya rastlanmamıştır. Bu eksikliğin giderilmesi ve bu alanda öncü bir çalışma olması için bu çalışma gerçekleştirilmiştir. Bu çalışma ile iş ilanlarının verimliliğini ve doğruluğunu artırmak, zengin içerik sağlamak ve daha doğal Türkçe iş ilanı oluşturulması sağlanmıştır. Üretken yapay zeka alanında yer alan doğal dil işlemede metin üretimi için Retrieval-Augmented Generation modeli kullanılmıştır. Çevrimiçi olarak ilan edilen iş ilanı verilerinden faydalanılarak anahtar kelime ve sözcük grupları kullanılıp istenilen iş ilanı metni oluşturulması hedeflenmiştir.

## II. GEÇMİŞ ÇALIŞMALAR

Retrieval-Augmented Generation ile metin üretiminde geçmişe dayalı veriler ve büyük dil modelleri kullanılarak yeni metinlerin üretilmesi sağlanır. Bu süreçte veriler vektör veri tabanı adı verilen bir depoda tutulur. Sorgu gönderildiğinde bu veri tabanından ilgili dokümanlar alınarak cevap metni üretilir. Büyük dil modelleri görev özelinde istenen sorgularda bazen yetersiz kalmaktadır. En önemli sorunlardan biri halüsinasyonlardır[3]. Bu problemde, büyük dil modeli akıcı ve doğru gibi görünen cevaplar üretebilir; ancak, bu yanıtlar gerçekte doğru bilgi içermeyebilir. Bundan dolayı metin içeriğini zenginleştirmek için dış kaynaktan gelen verilere ihtiyaç duyulmaktadır. Bu çalışmalarda dil modelini geliştirmek için harici veri kaynaklarından token alınması ile modelde daha az parametre kullanarak GPT-3 gibi meşhur dil modellerinden daha iyi performans gösterdiğini belirtmiştir[4]. Başka bir çalışmada Wikipedia gibi veri kaynaklarından elde edilen veriler, modelin ön eğitimi, ince ayarı ve çıkarım aşamalarında kullanılmaktadır. Denetimsiz öğrenme kullanılarak yapılan bu çalışmada zorlayıcı bir görev olan Open-QA’da ince ayar ve RAG ön eğitim model etkinliği gösterilmiştir[5]. Metin üretimi görevlerinde RAG oldukça yaygın olarak kullanılmaktadır. Literatürde doğal dil işleminin üretken yapay zeka alanında metin üretimi görevleri ile ilgili birçok çalışma yapılmıştır. Generation-Augmented Retrieval for Open-Domain Question Answering[2] çalışmasında İngilizce dilinde yapılan RAG ile getirilen sonucu daha iyi hale getirmek için girdi olarak verilen sorulara yanıtlar veya başlıklar eklenerek ilgili dokümanlarla eşleştirme işlemi daha verimli hale gelir. Sohbet botları insan sohbetini taklit ederek çalışan bilgisayar programlarıdır[9]. Bir başka metin üretme çalışması sohbet botları geliştirmek üzere yapılmıştır. Makine öğrenmesi modelleri kullanılarak yapılan sohbet botları önemli cevaplar üretse de büyük dil modeli kullanılarak yapılanlar daha insana benzer cevaplar üretmektedir. Vakayil ve arkadaşları[7] bu sebeple Llama-2 büyük dil modelini kullanarak RAG destekli bir sohbet botuyla çalışmalarını yüksek doğrulukla tamamlamışlardır. Vidiwell ve arkadaşlarının bu çalışmasında[6] RAG kullanılmış ve web kazıma yapılmıştır, dış kaynaktan gelen bu veriler ile büyük dil modeli, parametre verimli ince ayar tekniklerinden olan LoRA ve QLoRA gibi performansı artırıcı yöntemler kullanılarak ince ayar yapılmıştır. Aynı zamanda bu veriler vektör veri tabanına yerleştirilmiştir. İnce ayar yapılan büyük dil modeli kullanılarak RAG ile kullanıcıya daha doğru ve kapsamlı cevaplar verilmiştir. Literatürde ders asistanlığı yaparak öğrencilere faydalı olması için sohbet botları da geliştirilmiştir. Buna bir örnek de Leeds Üniversitesinde verilen veri madenciliği ve metin analitiği dersi için

LangChain ve RAG kullanılarak ders asistanlığı yapan sohbet botu yapılmıştır[8].

Literatürde, iş ilanlarıyla ilgili bilgisayar bilimleri ve yapay zeka teknikleri kullanılarak çeşitli çalışmalar gerçekleştirilmiştir. İşe alım sektörüne yönelik, konu modelleme tabanlı içerik odaklı bir iş öneri motoru geliştirilmiştir[10]. İşe alım sektöründe en uygun aday ve iş eşleştirmesi zorlayıcı ve zaman alıcı olarak görülmektedir. Bu çalışmada, yapılan içerik bazlı tavsiye çalışması ile kullanıcıya becerilerine ve niteliklerine göre en uygun iş ilanları önerilmiştir. Özellik vektörleri olarak konu modelleri ve n-gram torbası kullanılmıştır. Belirli bir iş kolu için istenilen becerilerin sıklığı TF-IDF algoritması kullanılarak hesaplanmıştır. Hibrid iş öneri sistemi yapan Toon De Pessemier ve arkadaşları[11] KNN ve içerik tabanlı algoritmaları birleştirerek hafif ve hızlı bir yaklaşım sunuyorlar. Adayın geçmiş iş ilanları etkileşimlerini dikkate alıp, bu iş ilanlarının özelliklerini eşleştirerek en benzer özelliklerdeki ilanları arar ve en uygun olan ilanları adaya önerir. Bir çalışmada[12] sistem iş ilanlarında geçen cümleleri üç ana kategoriye ayırmaktadır: teknik beceriler, temel beceriler ve bunlarla alakalı olmayan bilgiler. Verileri kümelendirirken diğer büyük dil modellerine göre daha kısa sürede ve daha az maliyetle işlemleri yaptığı için BERT modeli seçilmiştir. Bu model cümlelerin vektörlerini çıkarıp, anlamını analiz ederek doğru kategorilere ayrıştırılmasını sağlamıştır. İş ilanları arasındaki benzerlik, ilanlardan çıkarılan beceri vektörleri dikkate alınarak Öklid uzaklığı metriği kullanılarak ölçülmüştür. Akshay Bhola ve arkadaşları[13], büyük ölçekli bir devlet iş ilanı portalını analiz ettiklerinde iş ilanında olması gereken bazı becerilerin ve özelliklerin yazılmadığını, aslında iş ilanında eksik becerilerin olduğunu fark etmişler. Bu sorunun sebebi olarak da bu alanda daha az bilgisi olan insan kaynakları ve uzman işveren arasındaki iletişim boşluğu olabileceği tahmin edilmiştir. Eksik beceriler, Singapur devlet iş ilanı portalı kullanılarak yüzdelik olarak her iş ilanı için belirtilmiştir. Gereken becerilerin tahmin edilmesi BERT tabanlı bir model ile yapılmıştır. Bootstrapping yaklaşımı ile beraber sık tekrar eden birbirleriyle ilişkili kelimeler de elde edilmiştir.

## III. DENEYSEL ÇALIŞMALAR

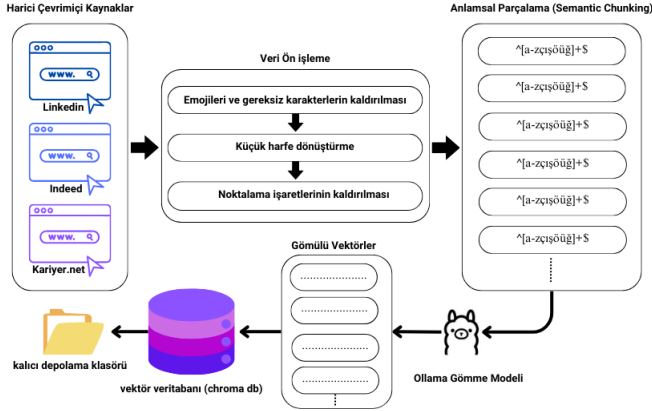
Bu bölümde veri setinin özellikleri, kullanılan modeller ve yöntemler, kütüphaneler, sonuçlar performans metriklerine göre sunulacaktır.

### A. Veri Seti ve Kullanılan Algoritmalar

Veri seti Türkçe yazılım iş ilanlarından oluşmaktadır. Bu iş ilanları çeşitli çevrimiçi web sitelerinden elde edilmiştir. Bu web siteleri LinkedIn, Kariyer.net ve Indeed’tir. Yapısal olmayan bu verilerin her biri iş ilanı başlığı ve iş tanımı olarak parçalara ayrılmıştır. Bu verilerin başlangıçlarına özellik isimleri eklenerek birleştirme yapıldı ve her bir iş ilanı tek bir veri haline dönüştü. Verilerde ön işlemede küçük harfe dönüştürme gerçekleştirilmiş ve alfabe, sayılar ve noktalama işaretlerinden bazıları silinmiştir. Nokta, artı ve sharp(#) gibi karakterler programlama dillerinin isimlerinde bulunduğu için silinmemiş ancak emojiler kaldırılmıştır. Tüm çalışmada Python framework’ü LangChain kullanılmıştır.

Elde edilen ön işlemden geçmiş veri seti anlamsal parçalama (semantic chunking) tekniği ile parçalara ayrılmıştır. Anlamsal parçalama probleme göre RAG

modellerindeki iyileştirmeleri ile son zamanlarda oldukça popüler hale gelmiştir[14]. Bu parçalama tekniğinde birbirleriyle ilişkili cümleler kümelenir ve parçalar oluşturulur. Sonrasında Ollama gömme modeli llama 3 kullanılarak bu parçalar vektör haline getirilmiştir. Bu gömülü vektörler Chroma vektör veri tabanına kaydedilmiştir.



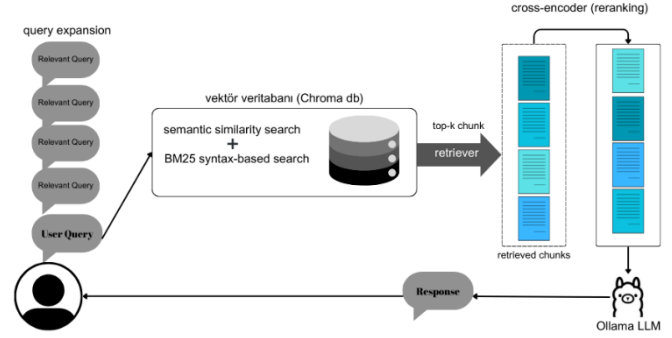
Şekil 1: Verilerin ön işlemden geçip vektör veritabanına kaydedilmesi

Bu çalışmada hibrid RAG modeli kullanılmıştır. Bu modelde hibrid getiriciler veri tabanını tarayarak en alakalı bilgileri getirmektedir. Hibrit getiriciler yoğunluk bazlı arama ve sözdizimi bazlı aramayı birleştirerek arama yapmaktadır. Bu çalışmada hibrid getiricilerin ağırlıkları yoğunluk bazlı getiricinin ve sözdizimi bazlı getiricinin sırasıyla 0.7 ve 0.3 olarak ayarlanmıştır. Bu ağırlıklar, çeşitli hiper-parametre ayarları üzerinde yapılan denemeler sonucunda optimize edilmiştir. Literatürde de yaygın olarak belirtilen benzer ağırlıkların kullanıldığı gözlemlenmiştir.

Veri tabanı hazırlandıktan sonra iş ilanına uygun bir soru oluşturulur. Bu soruya Groq API uygulanarak soru çoğaltma gerçekleştirilmiştir. Bu soru çoğaltma işleminde Ollama'nın llama-3.3-70b-versatile modeli kullanılmıştır. Oluşturulan dört farklı soru girdi olarak verilen soruya anlamsal olarak yakın olan ama farklı biçimde yazılmış olan sorulardır. Bu sorular ile vektör veri tabanından en alakalı dokümanlar elde edilmektedir. Burada doküman sayısı 300 olarak belirlenmiştir.

Doküman getirmelerinde soru ile dokümanlar arasında benzerlik ölçümü yapılmaktadır. Literatürde sözdizimi dikkate alınarak yapılan vektör veri tabanı aramaları BM25[15] ve TF-IDF[16] sıklıkla kullanılmaktadır. BM25 (Best Match) TF-IDF tabanlı bir yaklaşımdır. TF-IDF'in bazı dezavantajlarını azaltan daha verimli bir algoritmadır. İyileştirmelerin yapıldığı bazı alanlar şunlardır: doyunluk, doküman uzunluğu normalizasyonu, terim frekansı ağırlıklandırma. Bu çalışmada çok daha verimli ve literatürde de yaygın olarak kullanılan BM25 getirici kullanılmıştır. Getiriciler kullanılarak en iyi dokümanlar vektör veri tabanından çekilmiştir. Cross-Encoder kullanılarak dokümanlar tekrar sıralanmıştır. Cross-Encoder yöntemi Bi-Encoder yöntemine, eğer benzerlikleri ölçülecek metinlerin gömme vektörlerine ihtiyaç yoksa, tercih edilir. Corss-Encoder yöntemi ile daha doğru sonuçlar üretilir ve aynı zamanda bu yöntem daha verimlidir fakat daha yavaştır. Bu yöntem veri tabanından çekilen dokümanları yeniden sıralamak için literatürde kullanılmaktadır[17]. Bu çalışmada sıralamalarda soru ve dokümanlar arası kosinüs benzerlik metriği dikkate alınmıştır. Cross-Encoder tekniğinde kullanılan açık kaynak kodlu ve Türkçe dilini destekleyen

“roberta-base” kullanılmıştır. Bu model ile elde edilen gömme vektörleri arasındaki benzerlik kosinüs benzerliği ile ölçülmüştür. Daha sonra sıralanan dokümanlar büyük dil modeline verilerek sentezleme yapılır ve sonuç metni üretilmiş oluyor.



Şekil 2: Kullanıcı sorgusu ve cevap alma süreci

## B. Deneyisel Sonuçlar

Türkçe iş ilanı üretiminde girdi olarak verilen sorudan elde edilen çıktı ve gerçek veri arasındaki benzerlik ölçülmüştür. Bu benzerlik ölçümünde Levenshtein uzaklığı, Öklid uzaklığı, Jaccard benzerliği, TF-IDF tabanlı kosinüs benzerliği ve kosinüs benzerliği kullanılarak RoBERTa dil modeli ile anlam benzerliği ölçüm yöntemleri kullanılmıştır.

Levenshtein uzaklığı Levenshtein[Levenshtein (1966)] tarafından tanımlanan iki metin dizisi arasındaki farkı ölçen bir metriktir, bir diziyi diğerine dönüştürürken gereken değişim sayısını hesaplar. Sonuç değeri ne kadar düşük olursa, metinler arasındaki fark o kadar azdır ve dolayısıyla karşılaştırılan metinler birbirine o kadar yakın kabul edilir. Sözdizimsel benzerlik ölçümü için kullanılabilir.

Jaccard benzerliğinde iki metnin ortak kelimelerini alır bunu kesişim kümesi mantığı ile yapar sonra bu sayıyı her iki metindeki toplam eşsiz kelime sayısına böler. Ortaya çıkan sayı 0 ile 1 arasında bir sayıdır. Bu sayı 1'e yaklaştıkça metinler arası benzerlikleri artar. Öklid mesafe tabanlı metrikte iki veri arasındaki mesafe Öklid formülü ile hesaplanır, metinlerin temsili arasındaki mesafe ne kadar az ise benzerlik o kadar çoktur.

TF-IDF metriği bir kelimenin bir belgedeki frekansını ölçer. TF-IDF kosinüs benzerliği iki metnin TF-IDF matrislerindeki mesafe tabanlı benzerliklerini ölçmektedir.

RoBERTa dil modeli açık kaynak kodlu bir dil modelidir ve BERT tabanlıdır. BERT modelinden daha iyileştirilmiş, güçlü ve hızlı hale getirilmiş halidir ve Facebook AI Research tarafından 2019 yılında tanıtılmıştır[18]. Türkçe dilini desteklemesi açısından bu çalışmada bu dil modelinin kullanılması uygun görülmüştür. Benzerlik ölçümü yapılırken metinler önce parçalara ayrılır, sonra gömme vektörleri oluşturulur. Bu vektörler üzerinde kosinüs benzerliği uygulanır. Eğitim veri seti 1340 iş ilanından oluşmaktadır, her bir eğitim verisinde başlık ve içerik bilgisi bulunmaktadır. RAG ile eğitim veri seti 2414 alt dokümana bölünmüştür. 120 tane test verisinden elde edilen sonuçlar şu şekildedir:

TABLO I

TEST VERİSETİ ÜZERİNDE SONUÇLARIN SUNULMASI

Test Veriseti	Skorlar			
	Sadece Hibrid RAG	Soru Çoğaltmalı Hibrid RAG	Yeniden Sıralama Hibrid RAG	Büyük Dil Modeli
RoBERTa kosinüs benzerliği	0.994470	0.994405	0.994480	0.994365
TF-IDF kosinüs benzerliği	0.295081	0.297070	0.294003	0.277407
Jaccard Benzerliği	0.109414	0.109582	0.109738	0.100626
Levenshtein Uzaklığı	1431.314	1465.550	1446.032	1333.796
Öklid Uzaklığı	24.39793	24.83805	24.55276	22.218994

Her bir test verisi için RAG türlerinde 10'ar tane çıktı üretilmiştir. Sonra bu çıktıların ortalaması alınarak verinin başarısı hesaplanmıştır. Büyük dil modeli için de her bir test verisi için aynı şekilde 10 tane sonuç üretilmiş ve ortalaması alınmıştır. Daha stabil sonuç elde etmek için her bir test verisi için 10 tane cevap üretilmiştir. RAG'da metin üretirken kullanılan büyük dil modeli llama3.1:latest kullanılmıştır. Tabloda büyük dil modeli olarak kullanılan dil modeli de yine llama3.1:latest'tir. Bu skorlar her bir veri için kaydedilmiştir. En son tüm test veri setinin skorları kullanılarak ortalama sonuçlar üretilmiştir ve Tablo 1'de gösterilmiştir.

Modele soru çoğaltma ve yeniden sıralama eklenmesi ile başarı da artmaktadır. Soru çoğaltma aşamasından sonra yeniden sıralama yönteminin yapılması ile bir miktar iyileşme gözlemlenmiştir. Bu çalışmada daha kapsamlı olması, probleme uygunluğu ve daha çok anlama yönelik ölçüm yaptığı için RoBERTa model metriğinin sonuçları daha çok dikkate alınmıştır. Hibrid RAG modeli ortalama olarak %0.012 oranında kullanılan büyük dil modelinden daha başarılı sonuç vermiştir.

#### IV. SONUÇ

Türkçe iş ilanı metni üretmek için RAG modeli kullanılmıştır. Çeşitli çevrimiçi web sitelerinden elde edilen Türkçe iş ilanları ile anahtar kelimeler ile girilen sorgu ile benzer nitelikte iş ilanları üretilmiştir. Elde edilen sonuçlara bakıldığında RAG modeli büyük dil modelinden daha gelişmiş ve gerçek iş ilanlarına benzer nitelikte, metriklerin de belirttiği gibi, metinler üretmiştir. Ek olarak iş ilanı metninin çekiciliğini artırmak için “yeniliklere açık ve teknolojiyi yakından takip eden adaylarımıza kapılarımız açıktır.” gibi cümleler de ekleyerek tecrübeli ve tecrübesiz insan kaynaklarına yol gösterici metinler üretilmiştir. Literatüre doğal dil işleme alanında RAG tabanlı Türkçe iş ilanları üretme konusunda bir araştırma katkısı sağlanması hedeflenmiştir.

- [1] A. Vaswani et al., “Attention is all you need,” arXiv (Cornell University), Jan. 2017, doi: 10.48550/arxiv.1706.03762.
- [2] Y. Mao et al., “Generation-Augmented Retrieval for Open-domain question answering,” arXiv (Cornell University), Jan. 2020, doi: 10.48550/arxiv.2009.08553.
- [3] D. Dale, E. Voita, L. Barrault, and M. R. Costa-Jussà, “Detecting and mitigating hallucinations in machine translation: model internal workings alone do well, sentence similarity even better,” arXiv (Cornell University), Jan. 2022, doi: 10.48550/arxiv.2212.08597.
- [4] S. Borgeaud et al., “Improving language models by retrieving from trillions of tokens,” arXiv (Cornell University), Jan. 2021, doi: 10.48550/arxiv.2112.04426.
- [5] K. Guu, K. Lee, Z. Tung, P. Pasupat, and M.-W. Chang, “REALM: Retrieval-Augmented Language Model Pre-Training,” arXiv (Cornell University), Jan. 2020, doi: 10.48550/arxiv.2002.08909.
- [6] S. Vidiyelli, M. Ramachandran, and A. Dharunbalaji, “Efficiency-Driven Custom Chatbot development: unleashing LangChain, RAG, and Performance-Optimized LLM fusion,” Computers, Materials & Continua/Computers, Materials & Continua (Print), vol. 80, no. 2, pp. 2423–2442, Jan. 2024, doi: 10.32604/cmc.2024.054360.
- [7] S. Vakayil, D. S. Juliet, Anitha. J, and S. Vakayil, “RAG-Based LLM chatbot using Llama-2,” 2022 6th International Conference on Devices, Circuits and Systems (ICDCS), pp. 1–5, Apr. 2024, doi: 10.1109/icdcs59278.2024.10561020.
- [8] B. Alsafari, E. Atwell, A. Walker, and M. Callaghan, “Towards effective teaching assistants: From intent-based chatbots to LLM-powered teaching assistants,” Natural Language Processing Journal, vol. 8, p. 100101, Aug. 2024, doi: 10.1016/j.nlp.2024.100101.
- [9] L. K. Fryer, M. Ainley, A. Thompson, A. Gibson, and Z. Sherlock, “Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners,” Computers in Human Behavior, vol. 75, pp. 461–468, May 2017, doi: 10.1016/j.chb.2017.05.045.
- [10] S. Bansal, A. Srivastava, and A. Arora, “Topic Modeling Driven Content based Jobs recommendation Engine for recruitment industry,” Procedia Computer Science, vol. 122, pp. 865–872, Jan. 2017, doi: 10.1016/j.procs.2017.11.448.
- [11] T. De Pessemer, K. Vanhecke, and L. Martens, “A scalable, high-performance algorithm for hybrid job recommendations,” Proc. ACM Conf. Recommender Syst., pp. 1–4, 2016. doi: 10.1145/2987538.2987539.
- [12] R. Wang, Q. Chen, Y. Wang, L. Xiong, and B. Shen, “JobViz: Skill-driven visual exploration of job advertisements,” Visual Informatics, vol. 8, no. 3, pp. 18–28, Jul. 2024, doi: 10.1016/j.visinf.2024.07.001.
- [13] A. Bhola, K. Halder, A. Prasad, and M.-Y. Kan, “Retrieving Skills from Job Descriptions: A Language Model Based Extreme Multi-label Classification Framework,” Proceedings of the 17th International Conference on Computational Linguistics -, Jan. 2020, doi: 10.18653/v1/2020.coling-main.513.
- [14] R. Qu, R. Tu, and F. Bao, “Is semantic chunking worth the computational cost?,” arXiv (Cornell University), Oct. 2024, doi: 10.48550/arxiv.2410.13070.
- [15] S. Robertson and H. Zaragoza, “The Probabilistic Relevance Framework: BM25 and beyond,” Foundations and Trends® in Information Retrieval, vol. 3, no. 4, pp. 333–389, Jan. 2009, doi: 10.1561/15000000019.
- [16] A. Rajaraman and J. D. Ullman, Data Mining, Cambridge University Press, 2011, pp. 1–17.
- [17] Y. Li, F. Liu, I. Vulić, and A. Korhonen, “Improving Bilingual Lexicon Induction with Cross-Encoder Reranking,” arXiv (Cornell University), Jan. 2022, doi: 10.48550/arxiv.2210.16953.
- [18] Y. Liu et al., “ROBERTA: A robustly optimized BERT pretraining approach,” arXiv (Cornell University), Jan. 2019, doi: 10.48550/arxiv.1907.11692.