

# Yapay Zeka ve Makine Öğrenimi Araştırmaları için Akıllı Chatbot

BUSENUR GÖKLER

## 1. Giriş ve Motivasyon

### Projenin Amacı

Bu projedeki temel amacım, hızla büyüyen Yapay Zeka (AI) ve Makine Öğrenimi (ML) alanındaki güncel bilimsel makalelere, araştırmacılara ve kavramlara kolay, hızlı ve bağlamsal bir şekilde erişim sağlayan akıllı bir asistan geliştirmektir. Kullanıcıların doğal dilde sordukları sorulara, yapılandırılmış(structured) bir knowledge graph üzerinden anlamlı yanıtlar üretebilen bir chatbot sistemi oluşturmayı hedefleyerek projeme başladım.

### Karşılaştığım Problem ve Motivasyonum

AI ve ML alanları, her gün yüzlerce yeni makalenin yayınlandığı, sürekli gelişen ve karmaşıklaşan alanlardır. Bir araştırmacı, öğrenci veya bu alana merak duyan bir birey olarak, bu devasa bilgi yığınına takip etmek, ilgili çalışmaları bulmak, farklı kavramlar arasındaki ilişkileri anlamak ve güncel trendlerden haberdar olmak oldukça zorlayıcıdır. Geleneksel keyword-based arama motorları, genellikle spesifik ve derinlemesine query'ler için yüzeysel sonuçlar döndürmekte ya da alakasız sonuçlar üretebilmektedir.

Bu zorluklar, beni şu sorulara yöneltti:

- En son yayınlanan makalelere nasıl daha verimli ulaşabilirim?
- Belirli bir konsept üzerinde çalışan kilit araştırmacılar kimlerdir?
- Farklı AI teknikleri arasındaki bağlantılar nelerdir?
- Doğal dilimle soru sorup, gerçekten aradığım bilgiye hızlıca ulaşabilir miyim?

Motivasyonum, bu bilgi keşfi sürecini basitleştirmek, araştırmacılara ve öğrencilere zaman kazandırmak ve AI/ML literatürünün derinliklerinde kaybolmadan, hedefe yönelik bilgiye erişimi mümkün kılmak.

### Geliştirdiğim Çözüm Yaklaşımı

Bu problemlere çözüm olarak, Knowledge Graph tabanlı, LLM ile güçlendirilmiş ve RAG prensibini kullanan hibrit bir chatbot sistemi hayata geçireceğim. Sistemim, AI/ML makalelerini, yazarlarını, kurumlarını, kullanılan yöntemleri ve temel kavramları structured bir şekilde bir Neo4j graf database'de saklamakta ve kullanıcı sorgularına bu graph üzerinden yanıtlar üretmektedir.

### Projemden Beklediğim Faydalar

- Kullanıcıların karmaşık sorularına bile ilgili ve doğru yanıtlar sunmak.
- Literatür taraması ve bilgi toplama aşamalarını önemli ölçüde kısaltmak.

## 2. Veri Yükleme ve Knowledge Graph'ın Oluşturulması

Projem için gerekli AI/ML makalelerini simüle ederek CSV formatında veri setleri oluşturdum. Bu verileri, geliştirdiğim Cypher script'leri aracılığıyla Neo4j veritabanına nodes (Paper, Author, Concept vb.) ve relationships (AUTHORED, CITES vb.) şeklinde aktardım. Bu adım, chatbot'un üzerinde sorgu çalıştırabileceği zengin ve yapılandırılmış bir knowledge graph altyapısı sağladı.

### Node Türleri:

- **Paper:** {id, title, abstract, publication\_date, doi, venue, paper\_type, citation\_count, impact\_factor, keywords}
- **Author:** {id, name, h\_index, affiliation, research\_interests, orcid, country}
- **Concept:** {id, name, definition, category, synonyms, complexity\_level}
- **Method:** {id, name, description, category, computational\_complexity}
- **Institution:** {id, name, country, type, ranking, established\_year}
- **Year:** {value}

### Relationship Türleri:

- **AUTHORED:** (Author) → (Paper) {role, position}
- **AFFILIATED\_WITH:** (Author) → (Institution) {position, is\_current}
- **FOCUSES\_ON:** (Paper) → (Concept) {relevance\_score, is\_primary}
- **USES\_METHOD:** (Paper) → (Method) {usage\_type, performance\_improvement}
- **CITES:** (Paper) → (Paper) {citation\_context, relevance\_score}
- **RELATED\_TO:** (Concept) → (Concept) {relationship\_type, strength}
- **PUBLISHED\_IN:** (Paper) → (Year)
- **COLLABORATES\_WITH:** (Author) → (Author) {collaboration\_count, paper\_id}

Proje taslağı bu şekildedir, ancak projemin gidişatına göre değişiklikler (eklemeler, çıkartmalar) yapılabilir.

**Github ödev linki (Homeworks-5):** <https://github.com/Buse-ng/newmind-ai-bootcamp-learning/tree/main/homeworks>

**Github final proje linki:** <https://github.com/Buse-ng/NewMindAI-FinalCase>