

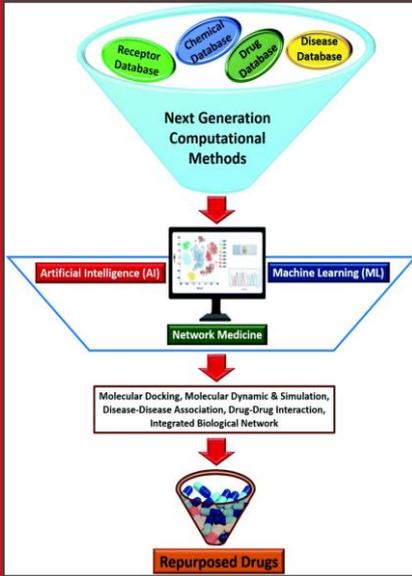


**Ahmed Marwan
Abdulhabeab
Alqershi**



ahmed.alqershi@agu.edu.tr

0009-0006-7886-1080



Thesis Advisor

**Asst. Prof.
Mehmet Gökhan
BAKAL**

gokhan.bakal@agu.edu.tr

30/05/2025

NEURAL INSIGHTS INTO DRUG REPOSITIONING: A LITERATURE-BASED FRAMEWORK USING WORD EMBEDDINGS AND SIAMESE NETWORKS

abstract The high cost, long timelines, and risks of traditional drug development have sparked interest in drug repositioning—finding new uses for drugs we already have. My thesis introduces a deep learning tool that digs into biomedical data from SemMedDB to spot potential new treatment connections between drugs and diseases. It's built to be a practical, efficient way to come up with ideas for early drug discovery.

The tool uses a Siamese Neural Network (SNN), trained on word patterns from the FastText model. The study tested two subnetworks—one dense, one convolutional—to see which worked best for pulling out useful features. After running over 570 setups, the best configuration hit a validation accuracy of 87.66% and a test accuracy of about 83%. It performed well across precision, recall, and F1-scores, too.

This work shows how deep learning paired with organized biomedical literature can power smarter drug discovery. It's not perfect yet—relying on one data source, using made-up negative samples, and missing contextual embeddings are some drawbacks. Overall, this system aims to support smarter decisions in drug research and fits into broader efforts for affordable and accessible healthcare.

keywords Drug Repositioning, Literature-Based Discovery, Siamese Neural Network, FastText Embeddings, Biomedical Text Mining

özet Geleneksel ilaç geliştirme süreçlerinin yüksek maliyetleri, uzun zaman çizelgeleri ve riskleri, mevcut ilaçların yeni kullanım alanlarını keşfetmeyi amaçlayan ilaç yeniden konumlandırma çalışmalarına olan ilgiyi artırmıştır. Bu tez, SemMedDB'den elde edilen biyomedikal verileri kullanarak, ilaçlar ile hastalıklar arasındaki potansiyel yeni tedavi bağlantılarını belirlemeye yönelik derin öğrenmeye dayalı bir sistem sunmaktadır. Geliştirilen sistem, erken aşama ilaç keşfi için pratik ve verimli bir fikir üretme yöntemi sağlamayı hedeflemektedir.

Sistem, FastText modelinden türetilen kelime desenlerini kullanarak eğitilen bir Siyam Sinir Ağı (SNN) mimarisine dayanmaktadır. Çalışmada, hangi yapının daha verimli özellikler çıkarabildiğini test etmek için biri yoğun (dense), diğeri evrişimli (convolutional) olan iki farklı alt ağ yapısı denenmiştir. 570'ün üzerinde model yapılandırması test edilmiş ve en iyi konfigürasyon %87.66 doğrulama doğruluğu ve yaklaşık %83 test doğruluğu elde etmiştir. Ayrıca kesinlik, duyarlılık ve F1-skorları açısından da dengeli bir performans sergilemiştir.

Bu çalışma, derin öğrenmenin organize edilmiş biyomedikal literatür ile birleşiminin, daha akıllı ilaç keşif süreçlerine nasıl katkı sağlayabileceğini göstermektedir.

anahtar kelime İlaç Yeniden Konumlandırma, Literatür Tabanlı Keşif, Siyam Sinir Ağı, FastText Gömülemeleri, Biyomedikal Metin Madenciliği