Interpretable Drug Repurposing Platform Based on Graph Neural Networks

Delta University for Science and Technology
جامعة الدلتا للعلوم والتكنولوجيا

Delta University for Science and Technology — Faculty of Artificial Intelligence — جامعة الدلتا للعلوم والتكنولوجيا — كلية الذكاء الاصطناعي

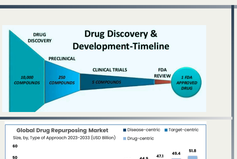
Motivation

Challenge

- 1 in 10 affected by rare diseases
- 95% lack treatments
- 3–15 yrs to diagnose, 3–5× higher costs
- Drug discovery: 15 yrs, 10,000 → 1 approval

Opportunity

- Repurposing = faster, cheaper, safer
- Market growing to $51.8B by 2033
- Driven by data, AI & GNNs



Rare Diseases by the Numbers

- **1 in 10 people** are affected by a rare disease.
- **1 in 2 people** diagnosed with a rare disease are children.
- **3–15 years** is a common timeline for diagnosis.
- **95% of rare diseases** lack an FDA-approved treatment.
- People with rare diseases face **3–5 times higher medical costs** than people with non-rare diseases.

Objective

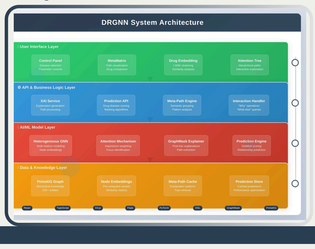
Accelerate Drug Repurposing
Using GNNs to identify novel drug-disease links efficiently.

Enable Interpretability
Integrate explainable AI (GraphMask, attention) to reveal why predictions are made.

Deliver Interactive Insights
Through a user-friendly UI (MetaMatrix, Attention Tree) for drug/path visualization.

Leverage Biomedical Knowledge
Utilize PrimeKG and meta-path reasoning for deeper biomedical understanding.

Build Scalable, Modular System
Full-stack architecture (React, Flask, PyTorch) for seamless integration and performance.



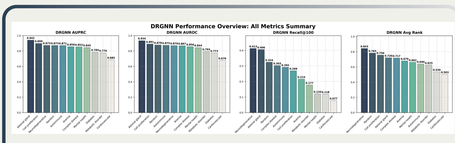
Results

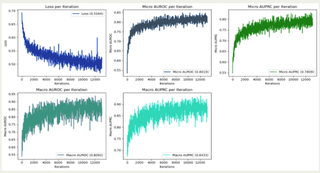
Overall Performance

DRGNN achieves exceptional performance across all evaluation metrics:

Key Metrics

AUPRC: 0.942 (excellent precision-recall balance)
AUROC: 0.934 (strong discriminative ability)
Average Rank: 0.843 (superior ranking performance)
Recall@100: 0.413 (good retrieval coverage)



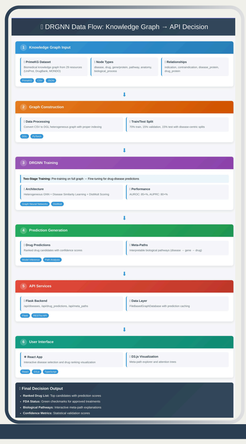


Team Members

- Salah Gamal Abdelkhabir
- Ahmed Ebrahim Gabr
- Ahmed Hamdy Eldakroury
- Ahmed Mohamed Ali
- Sara Ayman Ibrahim
- Aliaa Mohamed Ali

Materials & Methods

- **Knowledge Graph Input**
  Utilizes PrimeKG with rich biomedical entities (drugs, diseases, genes) and relations.
- **Graph Construction**
  Preprocessed into a DGL-compatible graph with disease-centric train/test split.
- **Model Training**
  Two-stage process: pre-training on the full graph and fine-tuning for drug-disease links using a heterogeneous GNN with DistMult scoring.
- **Prediction Generation**
  Outputs ranked drug candidates and interpretable meta-paths (disease → gene → drug).
- **API Services**
  Flask-based REST API delivers predictions and explanation paths.
- **User Interface**
  React + D3.js frontend enables disease selection, drug ranking, and pathway visualization.
- **Final Output:**
  Ranked drug list, FDA status, meta-path explanations, and confidence metrics.



DRGNN Platform

- **Interactive Disease Selection**
  Users input a disease (e.g., Macroglobulinemia) to generate tailored drug predictions.
- **Ranked Drug Outputsa**
  Top candidates are scored for effectiveness, such as Omega Interferon and Metoprine (score ≈ 1.000).
- **Drug Embedding Visualization**
  A 2D embedding map shows drug clusters, highlighting predicted treatments in green for quick insight.
- **Meta-Path Explanations**
  Visual pathways (e.g., disease → gene → anatomy → drug) uncover mechanistic links behind each prediction, enhancing trust and interpretability.
- **Path Filtering & Exploration**
  Users can filter by node type, edge score, or pathway structure to refine and understand model decisions.

Future Work

- **Expanded Disease Coverage:** Integrate more diseases across therapeutic areas.
- **Multi-Modal Data Fusion:** Incorporate omics, clinical trials, and literature data.
- **Advanced Interpretability:** Improve meta-path clarity using natural language and causal reasoning.
- **Real-World Validation:** Collaborate with biomedical labs for experimental verification.
- **Continuous Learning:** Enable model updates with new biomedical knowledge in real time.

Conclusion

DRGNN is a robust and interpretable platform for AI-driven drug repurposing. By integrating knowledge graphs, GNNs, and explainable meta-paths, it delivers accurate predictions and meaningful insights. With a transparent, end-to-end pipeline, DRGNN supports research, clinical workflows, and decision-making—positioning it as a key tool in the future of precision medicine.