

# DeepGLSTM: Deep Graph Convolutional Network and LSTM based approach for predicting drug-target binding affinity

Shrimon Mukherjee<sup>12</sup> Madhusudan Ghosh<sup>12</sup>  
Partha Basuchowdhuri<sup>1</sup>

<sup>1</sup>School of Mathematical and Computational Sciences  
Indian Association for the Cultivation of Science

<sup>2</sup>Equal Contribution to this work

SIAM International Conference on Data Mining, April 2022



# Table of Contents

- 1 Introduction
- 2 Proposed Model
- 3 Result
- 4 Case Studies
- 5 Analysis and Discussions
- 6 Conclusion

# Table of Contents

1 Introduction

2 Proposed Model

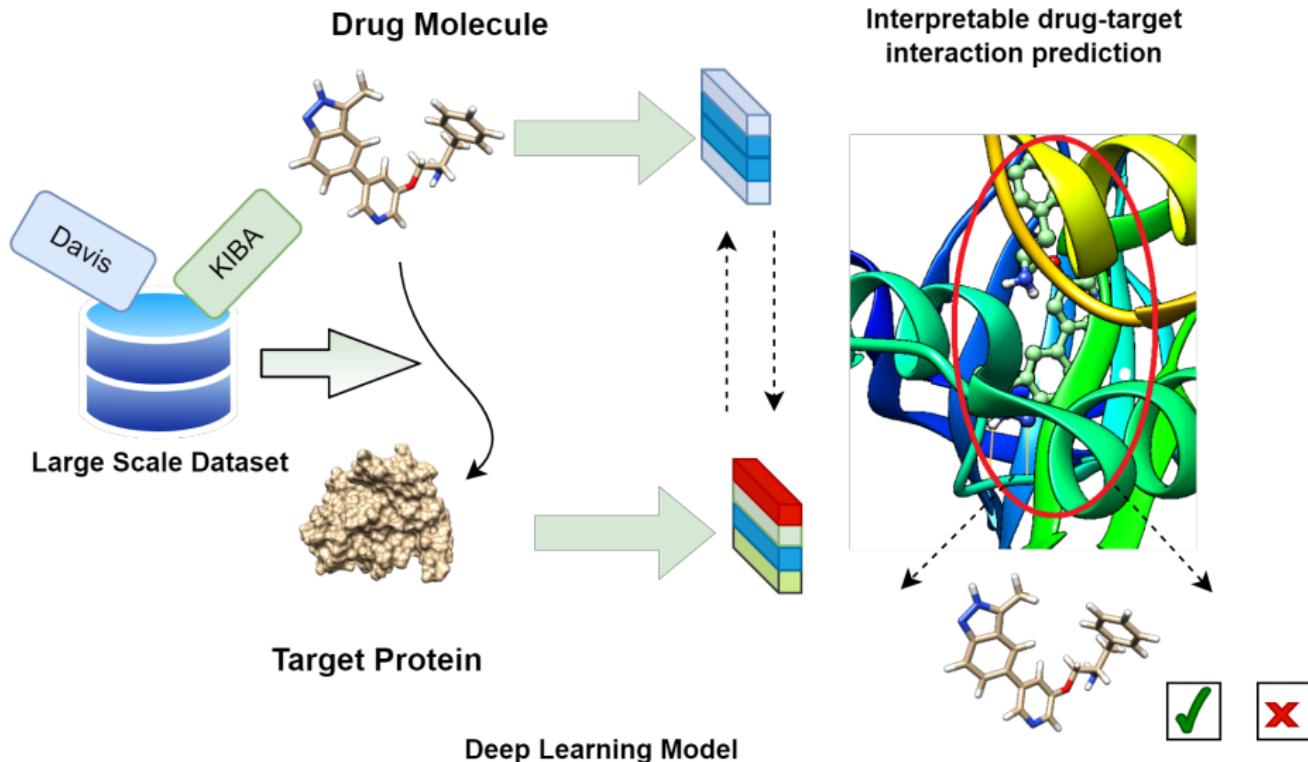
3 Result

4 Case Studies

5 Analysis and Discussions

6 Conclusion

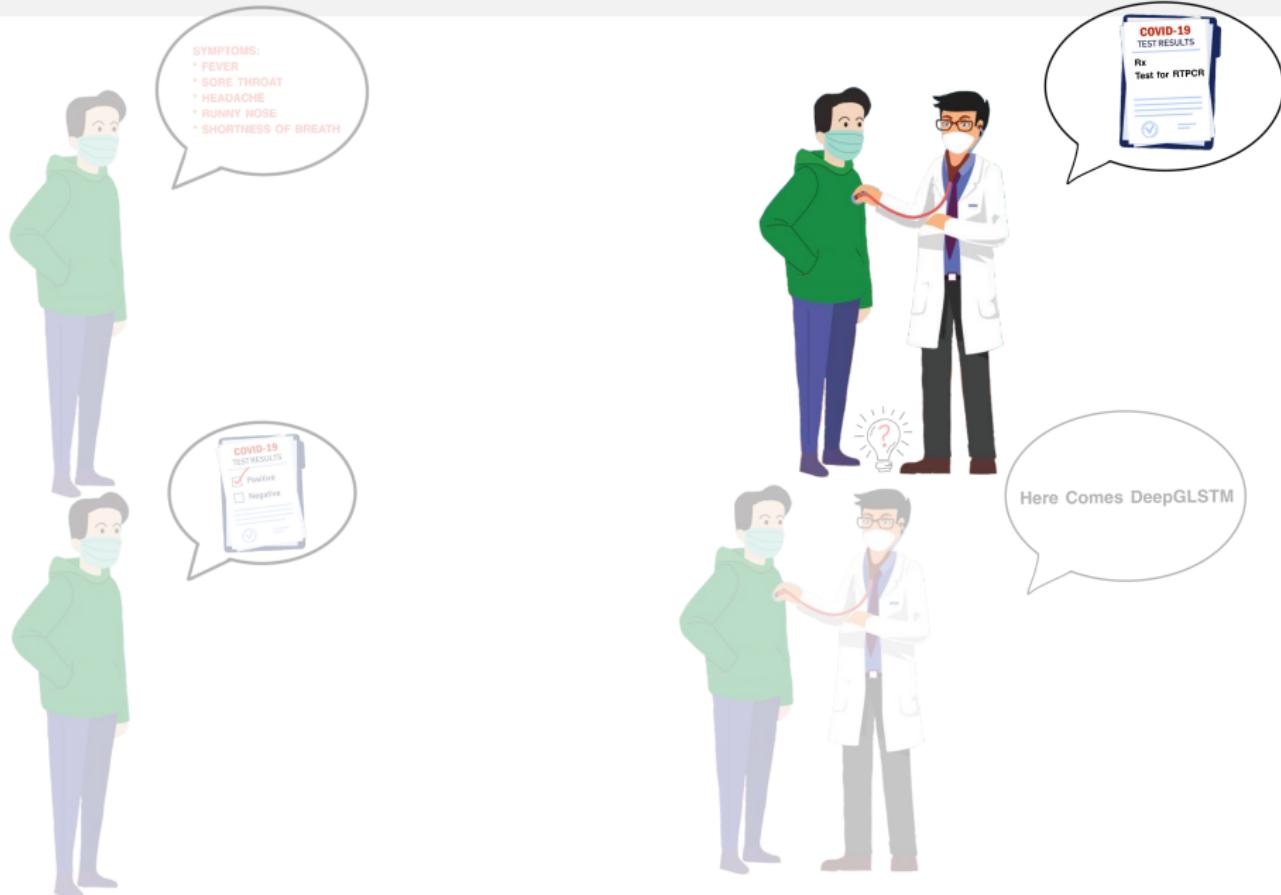
# Drug-target binding affinity prediction: Problem definition



# Importance of DeepGLSTM on current scenario:



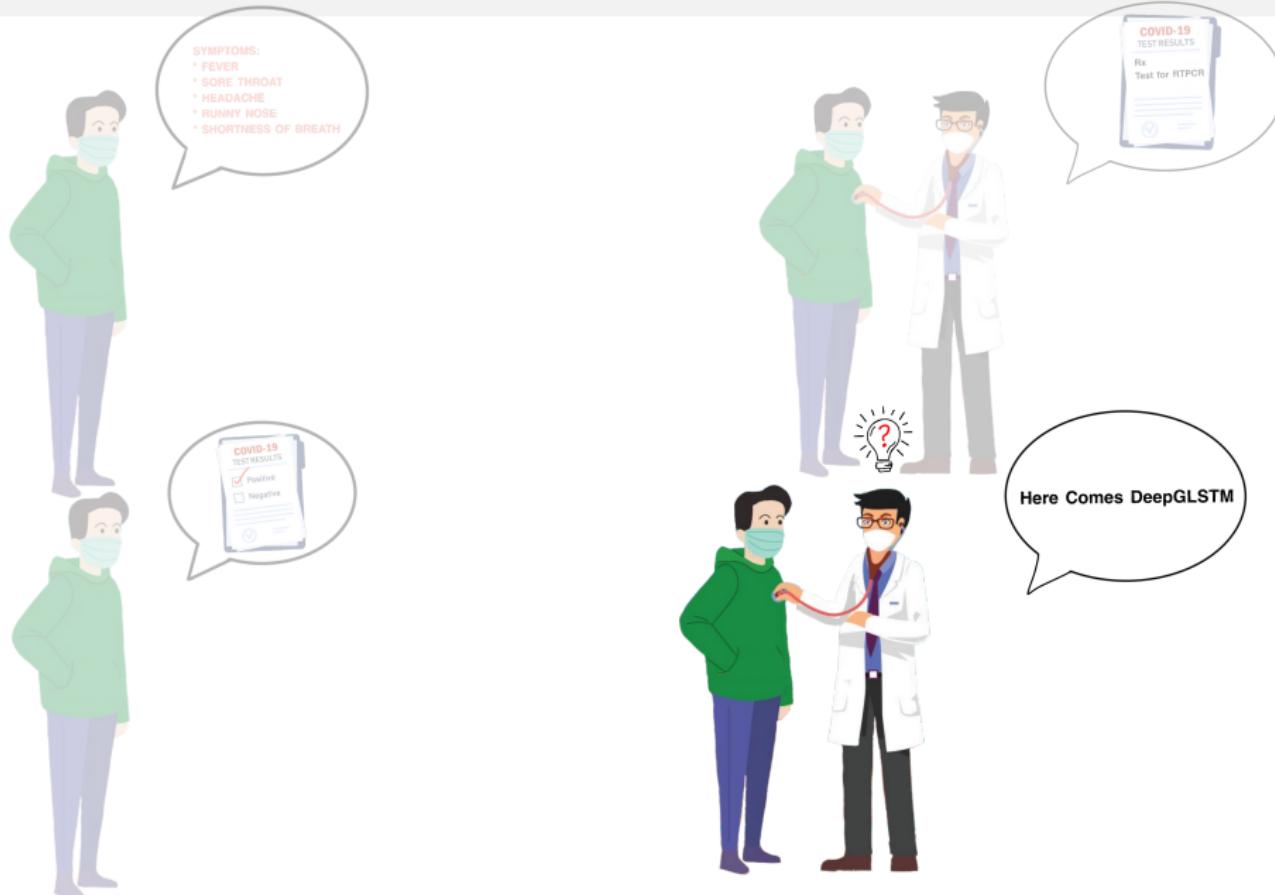
# Importance of DeepGLSTM on current scenario:



# Importance of DeepGLSTM on current scenario:



# Importance of DeepGLSTM on current scenario:



# Table of Contents

1 Introduction

2 Proposed Model

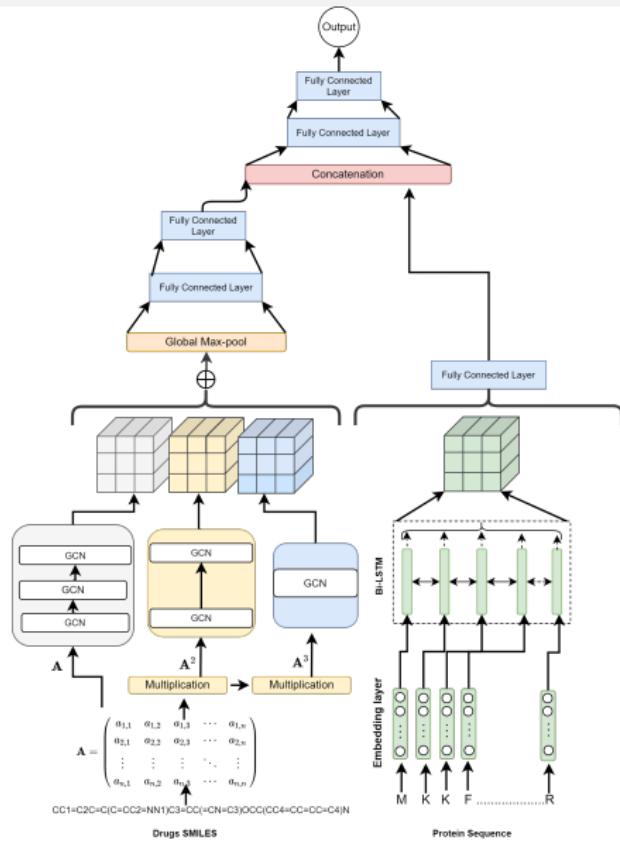
3 Result

4 Case Studies

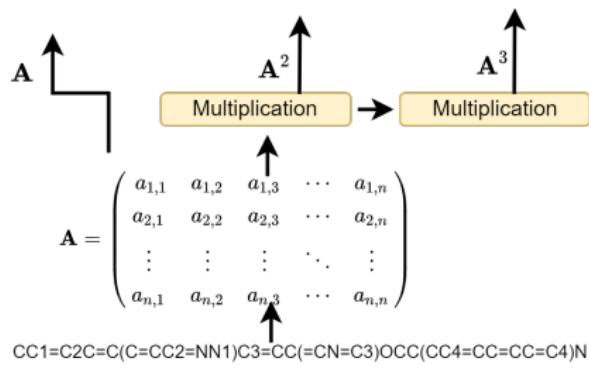
5 Analysis and Discussions

6 Conclusion

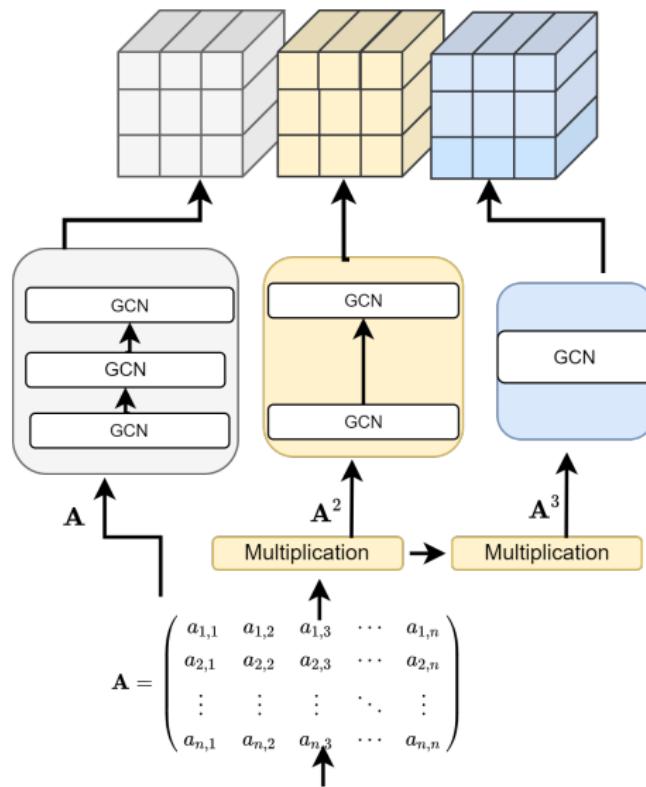
# DeepGLSTM - Complete Architecture



# GCN input block description

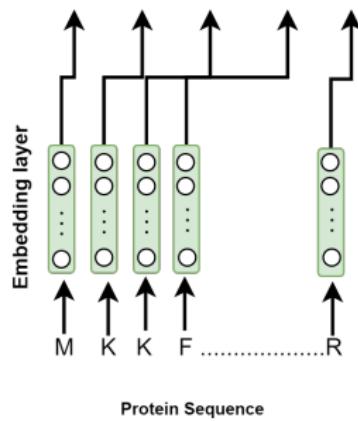


# GCN block description

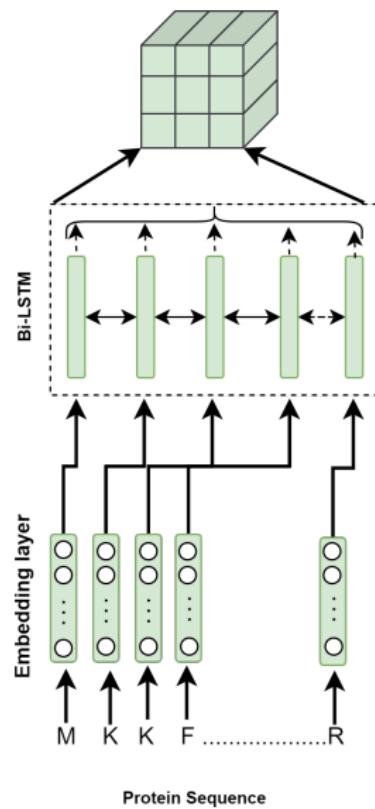


Drugs SMILES

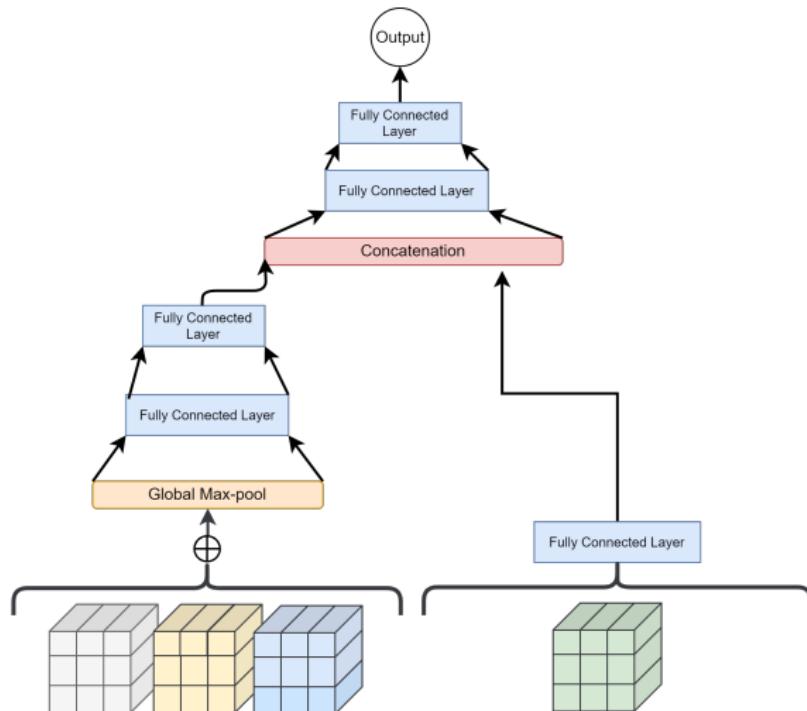
# Bi-LSTM block input description



# Bi-LSTM block description



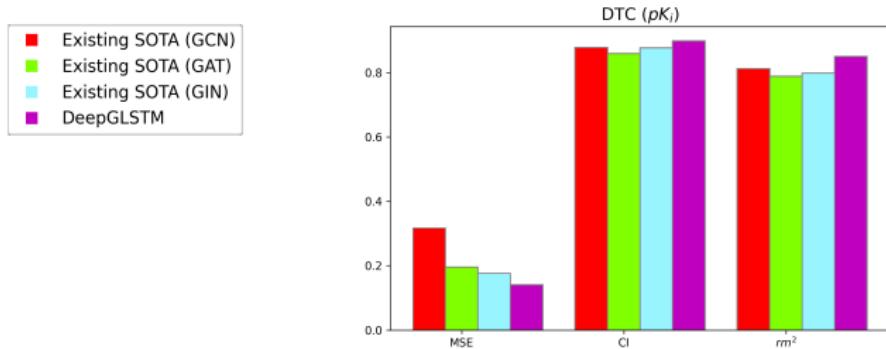
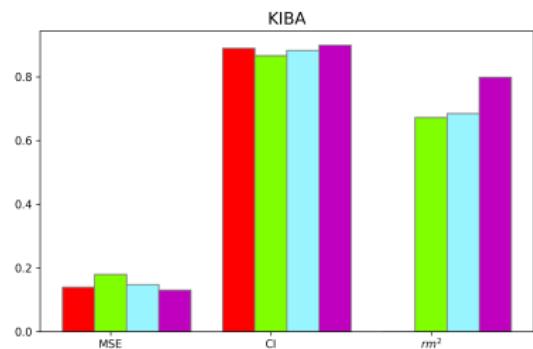
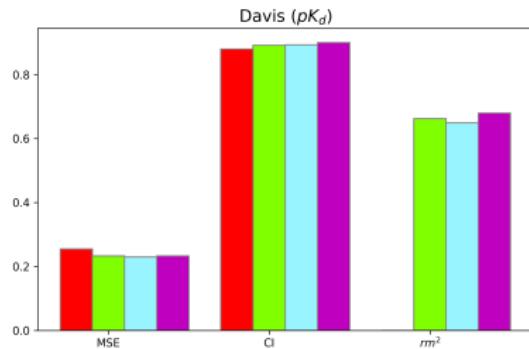
# Final block description



# Table of Contents

- 1 Introduction
- 2 Proposed Model
- 3 Result
- 4 Case Studies
- 5 Analysis and Discussions
- 6 Conclusion

# DeepGLSTM vs existing state-of-the-art

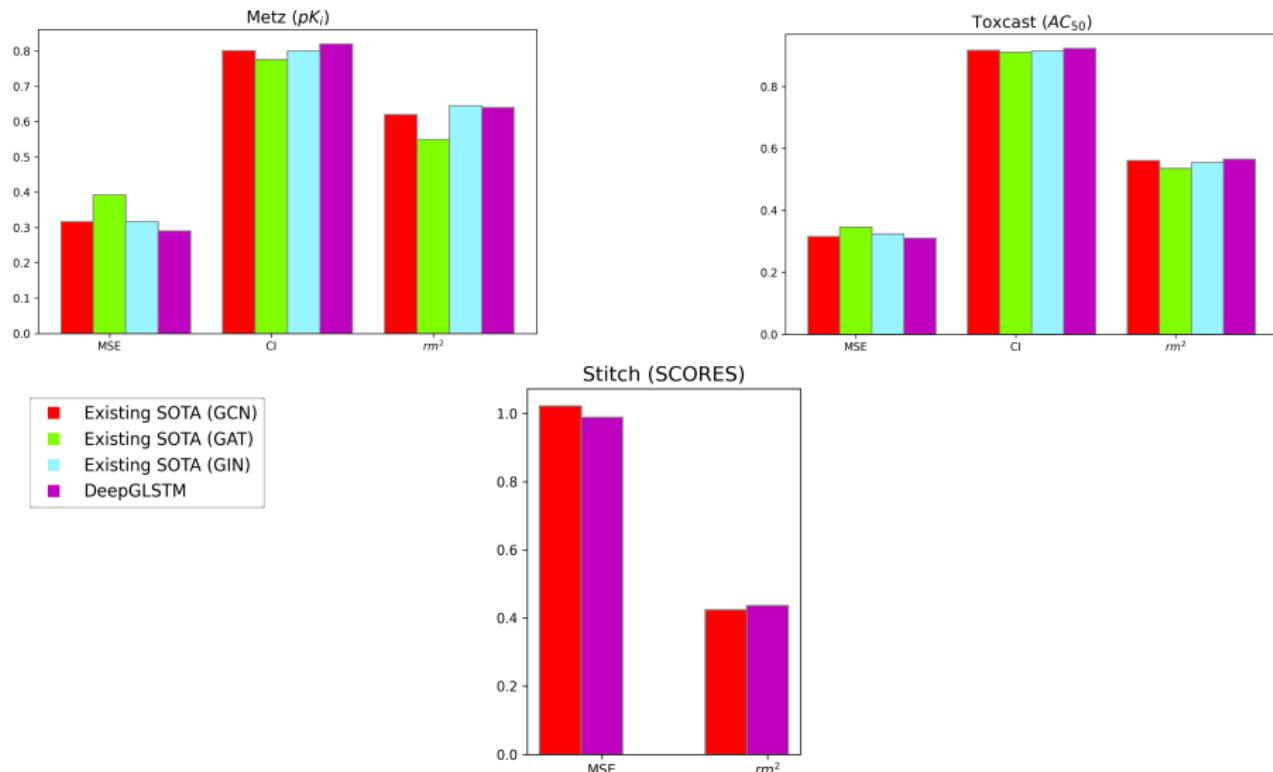


Legend:

- Existing SOTA (GCN)
- Existing SOTA (GAT)
- Existing SOTA (GIN)
- DeepGLSTM

Metric Scores (Y-axis) for Davis ( $pK_d$ ), KIBA and DTC ( $pK_i$ )

# DeepGLSTM vs existing state-of-the-art

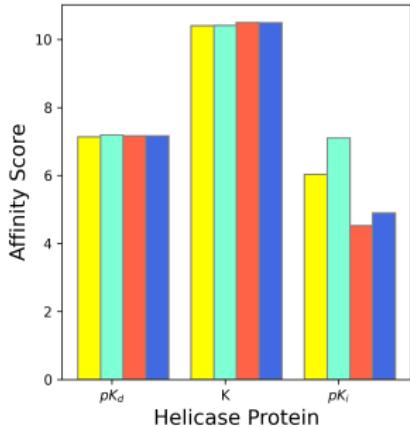
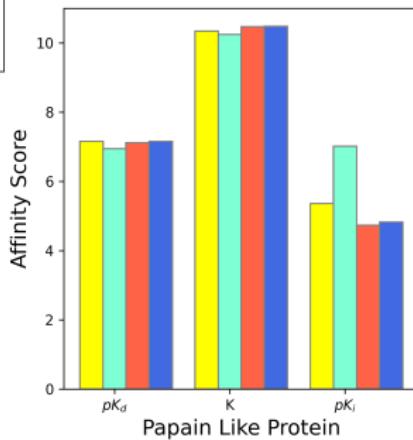
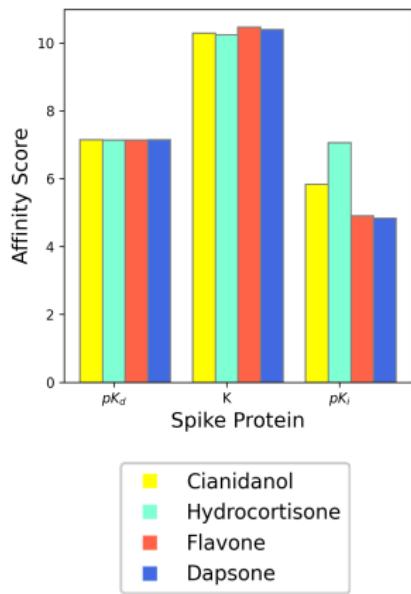


Metric Scores (Y-axis) for Metz ( $pK_i$ ), Toxcast ( $AC_{50}$ ) and Stitch (SCORES)

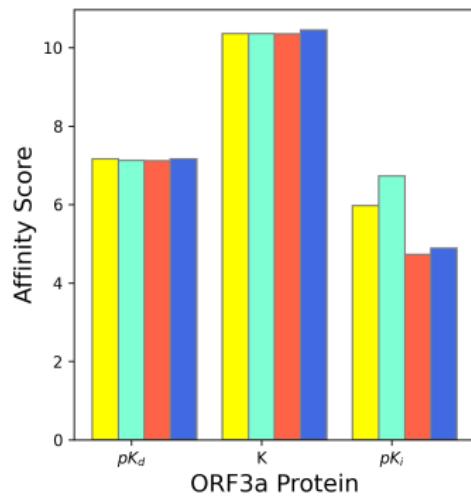
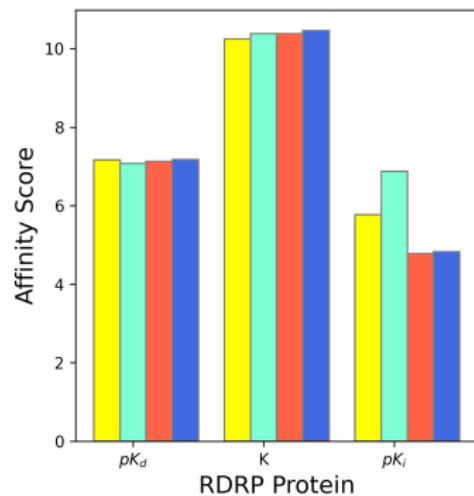
# Table of Contents

- 1 Introduction
- 2 Proposed Model
- 3 Result
- 4 Case Studies
- 5 Analysis and Discussions
- 6 Conclusion

# Case studies on SARS-CoV-2 viral proteins



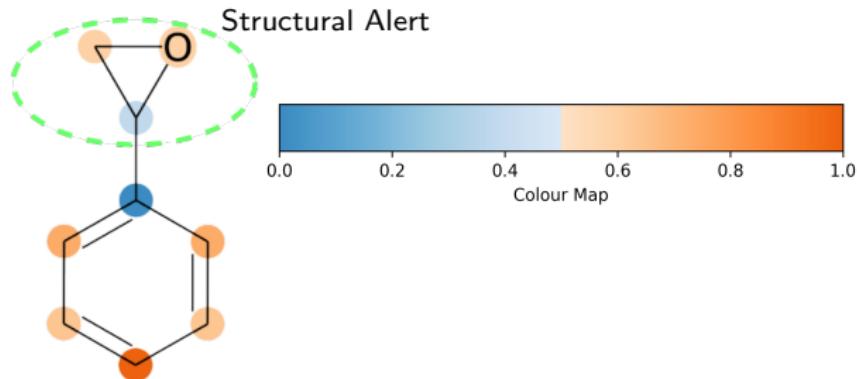
# Case studies on SARS-CoV-2 viral proteins



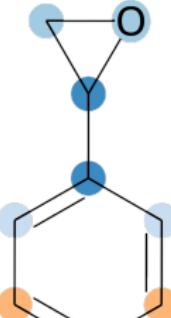
- Cianidanol
- Hydrocortisone
- Flavone
- Dapsone

# Finding structural alerts from chemical compounds

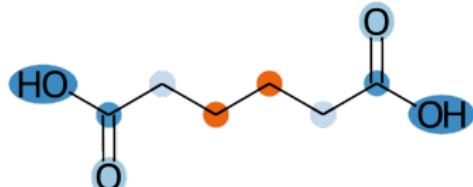
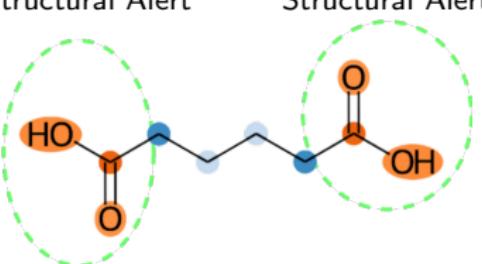
DeepGLSTM



GraphDTA (GAT)



Structural Alert

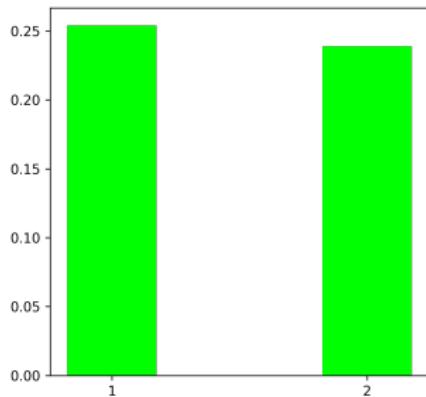


# Table of Contents

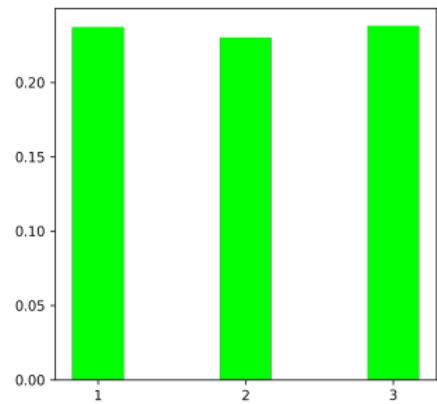
- 1 Introduction
- 2 Proposed Model
- 3 Result
- 4 Case Studies
- 5 Analysis and Discussions
- 6 Conclusion

# Analysis and Discussions

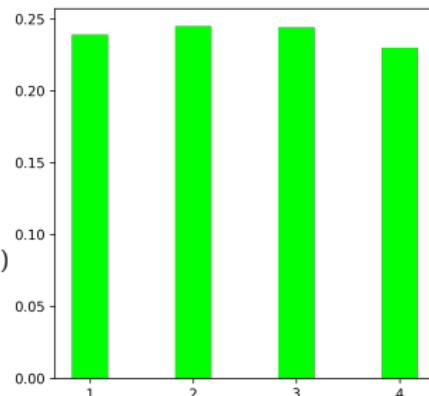
Performance comparison of 1-D CNN vs. Bi-LSTM



Effectiveness of different components of DeepGLSTM



Effectiveness of using the power graph



MSE Score (Y-axis) for Davis ( $pK_i$ )

# Table of Contents

1 Introduction

2 Proposed Model

3 Result

4 Case Studies

5 Analysis and Discussions

6 Conclusion

# Conclusion

- We present a novel architecture for the task of predicting binding affinity values between FDA-approved drugs and viral proteins. Empirical results show that our method is more accurate than the present state-of-the-art methods.
- Our model uses three blocks of GCN for learning the topological information of the drug molecules.
- Bi-LSTM component of our model learns the representation of the protein sequences.
- We also use our model for predicting binding affinity values between FDA-approved drugs and 5 viral proteins for SARS-CoV-2 to reveal prospective drugs that could be effective against them.

## Reference I

-  Cheng, F. In silico oncology drug repositioning and polypharmacology. *Cancer Bioinformatics*. pp. 243-261 (2019)
-  Nguyen, T., Le, H., Quinn, T., Nguyen, T., Le, T. & Venkatesh, S. GraphDTA: Predicting drug–target binding affinity with graph neural networks. *Bioinformatics*. **37**, 1140-1147 (2021)
-  Kipf, T. & Welling, M. Semi-Supervised Classification with Graph Convolutional Networks. *5th International Conference On Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. (2017),  
<https://openreview.net/forum?id=SJU4ayYgI>
-  Davis, M., Hunt, J., Herrgard, S., Ciceri, P., Wodicka, L., Pallares, G., Hocker, M., Treiber, D. & Zarrinkar, P. Comprehensive analysis of kinase inhibitor selectivity. *Nature Biotechnology*. **29**, 1046-1051 (2011)

## Reference II

-  Tang, J., Szwajda, A., Shakyawar, S., Xu, T., Hintsanen, P., Wennerberg, K. & Aittokallio, T. Making sense of large-scale kinase inhibitor bioactivity data sets: a comparative and integrative analysis. *Journal Of Chemical Information And Modeling.* **54**, 735-743 (2014)
-  Landrum, G. Rdkit documentation. *Release.* **1**, 4 (2013)
-  Fey, M. & Lenssen, J. Fast Graph Representation Learning with PyTorch Geometric. *ICLR Workshop On Representation Learning On Graphs And Manifolds.* (2019)



Shrimon Mukherjee



Madhusudan Ghosh



Partha Basuchowdhuri

*Thank You!*

<https://arxiv.org/pdf/2201.06872.pdf>

<https://github.com/MLlab4CS/DeepGLSTM.git>

