

Long short-term memory neural network for drug-target interaction prediction

Konstantin Kalitin^{1,2}, Olga Mukha^{1*} and Alexander Spasov^{1,2}

¹ Volgograd State Medical University, 1, Pavshikh Bortsov Sq., Volgograd, 400131, Russian Federation

² Volgograd Medical Research Center, 1, Pavshikh Bortsov Sq., Volgograd, 400131, Russian Federation

* Correspondence: olay.myha14@gmail.com

Introduction

Modern drug discovery primarily concentrates on identifying and understanding drug-target interactions. Traditional techniques, limited by factors like throughput, precision, and cost, struggle to efficiently identify these potential drug-target interactions. Hence, there's a pressing need for advanced computational methods to address this challenge.

Materials and Methods

We constructed a deep learning model for drug-target interaction prediction. The features of target proteins (using iFeature Web Server) were extracted and associated with drug molecular substructure fingerprints (PubChem Substructure Fingerprint) to form 3618--dimensional feature vectors of drug-target pairs. The features of compounds and target proteins were subsequently compressed into a unified vector space using sparse principal component analysis. Finally, we used a long short-term memory (LSTM) neural network to make predictions. Five-fold cross-validation was employed to evaluate the performance of our model.

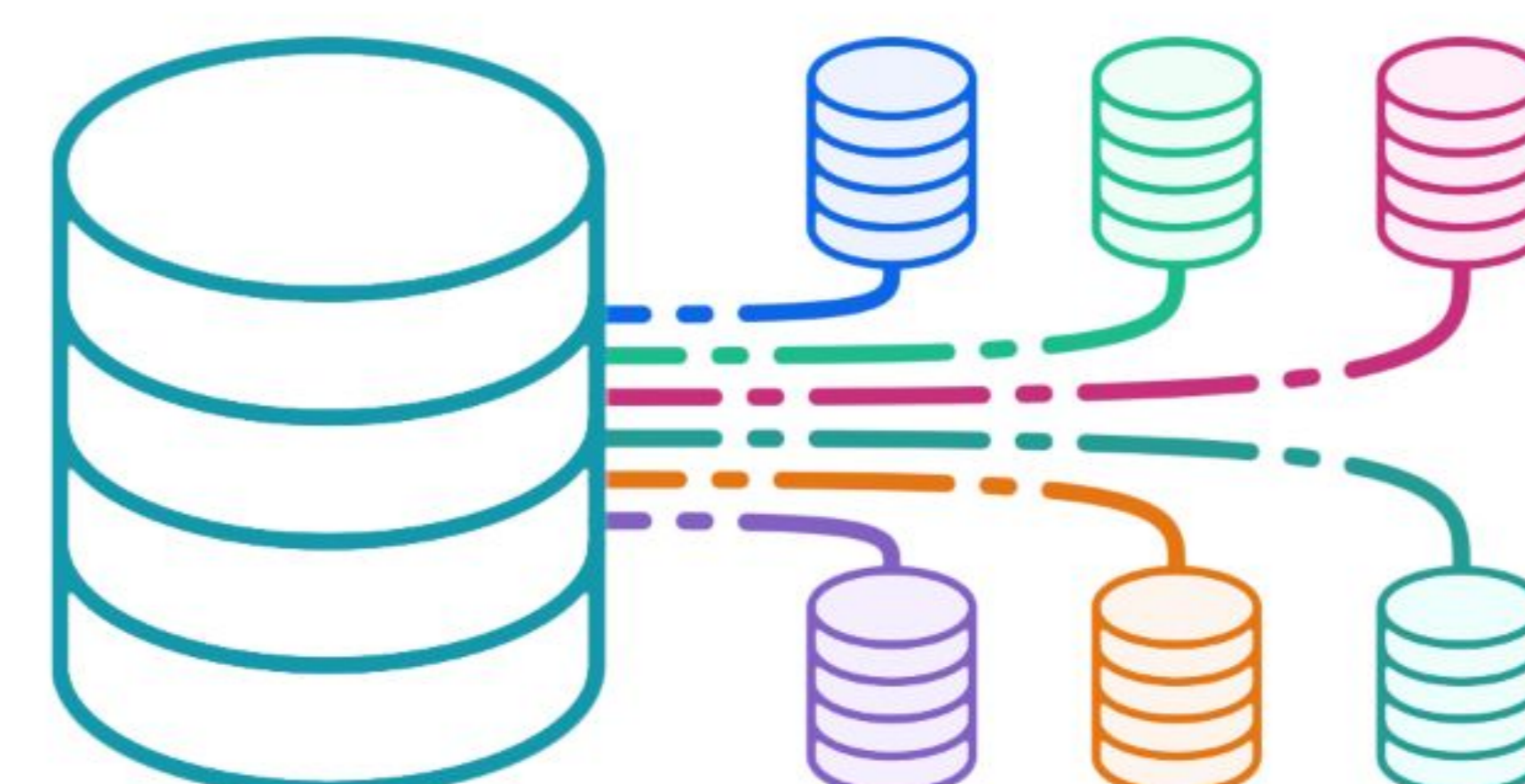
Results

Upon evaluation, our model showcased satisfactory performance in drug-target interaction prediction. Specifically, it achieved accuracies of 86.7%, 84.7%, and 73.4%. These scores were obtained from three different drug-target datasets, highlighting the model's robustness and generalizability. The slight variation in accuracy scores across datasets suggests that, while the model is highly effective, there might still be room for further optimization, particularly for datasets with unique characteristics.

Conclusion

These findings indicate that the method is competitive with other contemporary drug-target prediction tools.

Collecting DTI data from Public Databases



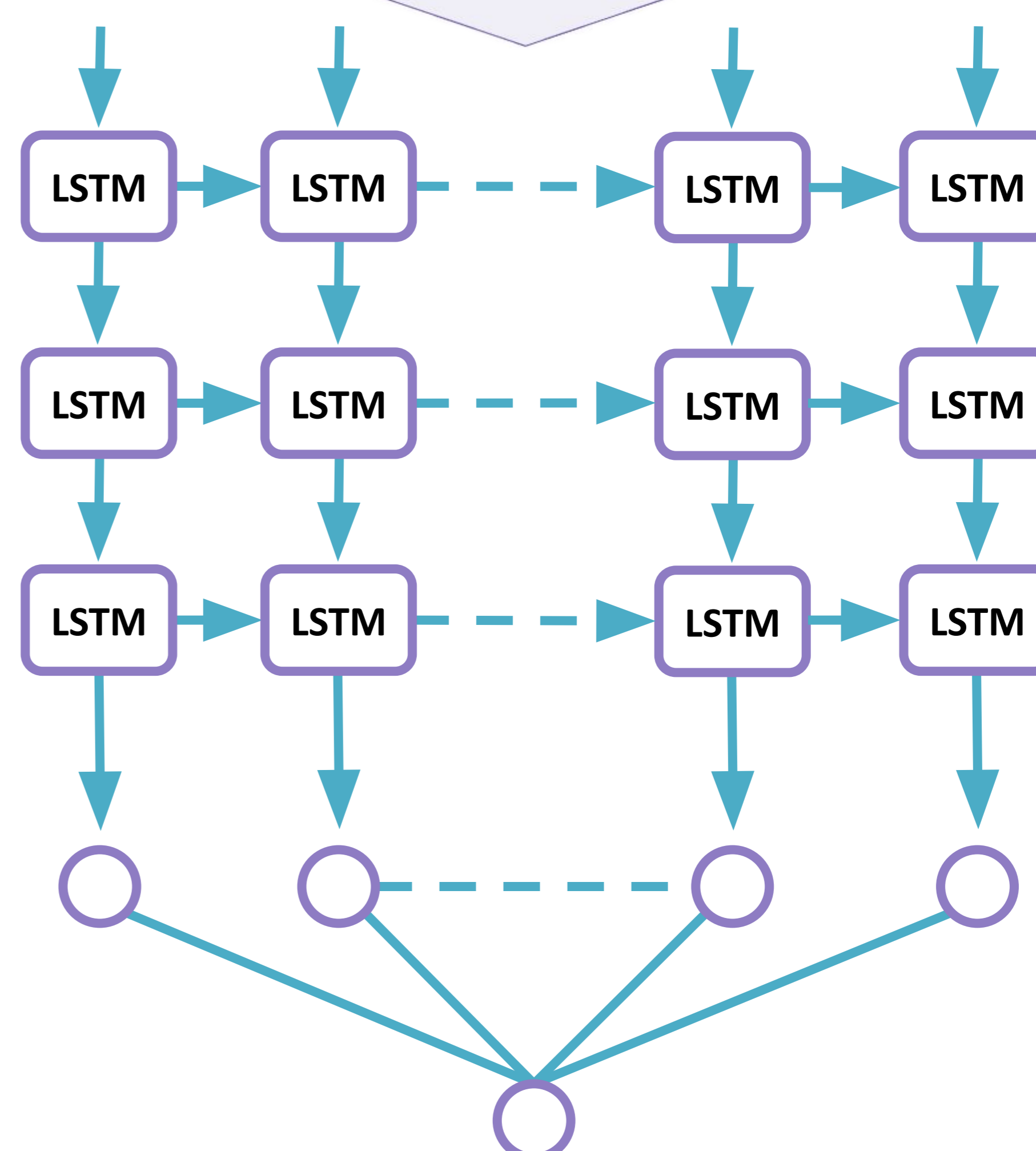
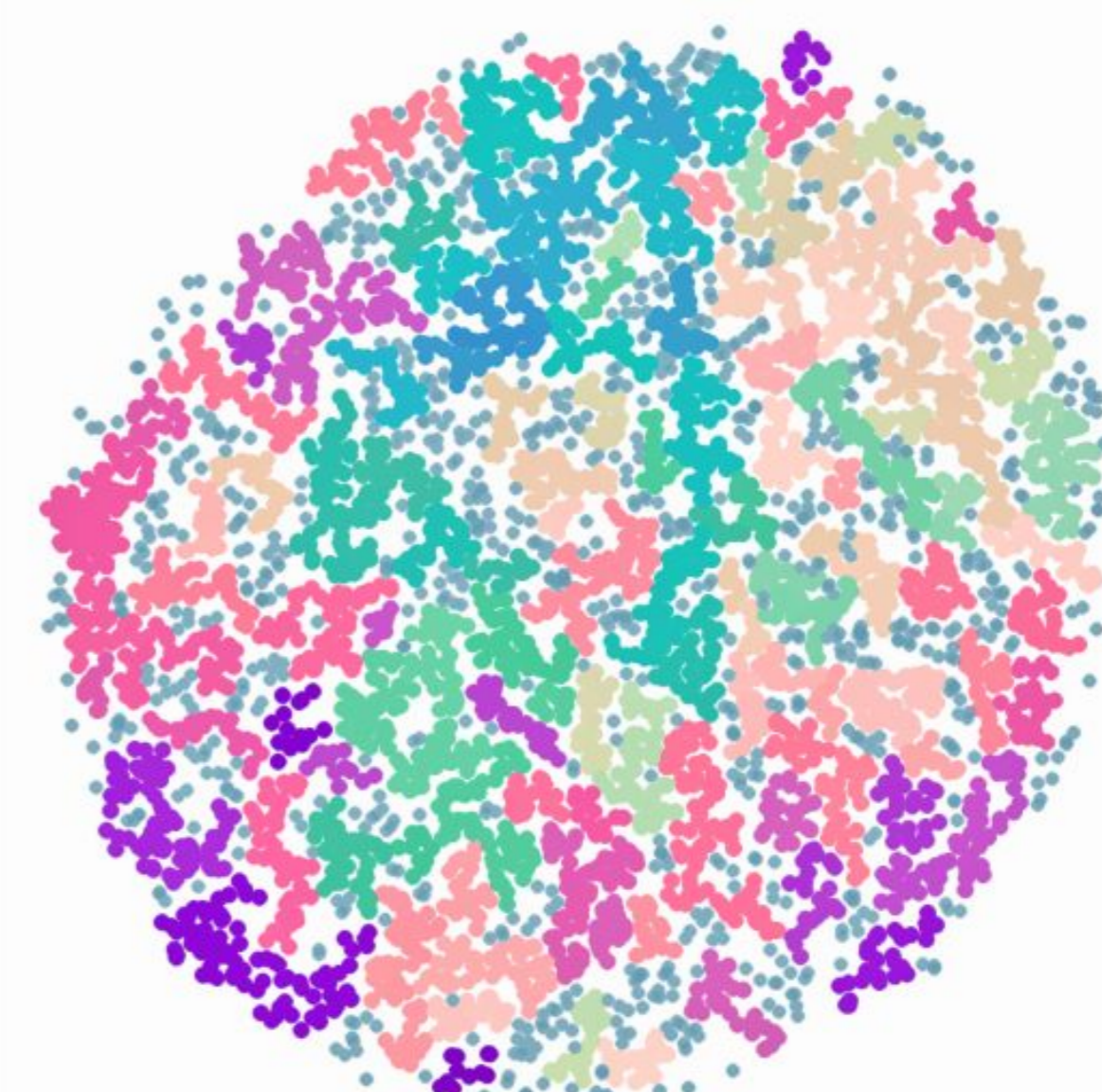
Drug Encoding

1	1	0	0	0
1	1	0	0	0
0	0	1	0	1
1	0	0	0	0
1	1	1	1	1

Protein Feature Extraction

1	1	0	0	0
1	1	0	0	0
0	0	1	0	1
1	0	0	0	0
1	1	1	1	1

Sparse Principal Component Analysis



The 9th International Electronic Conference on
Medicinal Chemistry

01–30 November 2023 | Online

