

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ
Кафедра биомедицинской информатики

**Разработка глубокой нейронной сети для оценки энергии
связывания химических соединений с молекулярной мишенью**

Курсовая работа

Благодарного Артёма Андреевича
ФПМИ, БМИ, 3 курс 3 группа

Научный руководитель:
профессор, доктор
физико-математических наук,
Тузиков А.В.

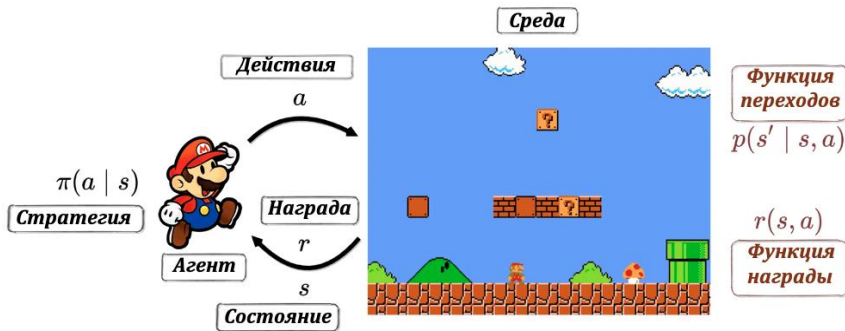
Цель работы :

Разработать глубокую нейронную сеть для оценки энергии связывания химических соединений с молекулярной мишенью

Задачи работы:

1. Изучить теорию по алгоритмам обучения с подкреплением.
2. Подобрать данные для обучения.
3. Разработать и обучить нейронную сеть.
4. Проанализировать результат.

Обучение с подкреплением



Докинг

Докинг — это метод, применяемый в структурной биоинформатике и хемоинформатике, с целью моделирования взаимодействия между двумя молекулами: как правило, небольшой молекулы-лиганда и более крупной мишени, чаще всего — белка. Задача докинга заключается в том, чтобы предсказать наиболее вероятную пространственную конфигурацию комплекса «белок–лиганд» и оценить силу этого взаимодействия.

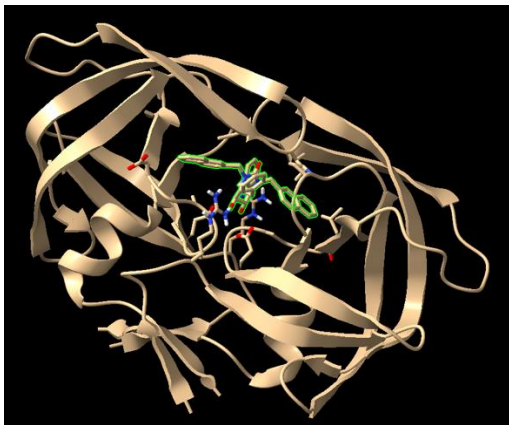


Рис. HIV-1 протераза с выделенным лигандом

Данные для обучения LP-PDBBind

THGLab Teresa Head-
Gordon's Research Group

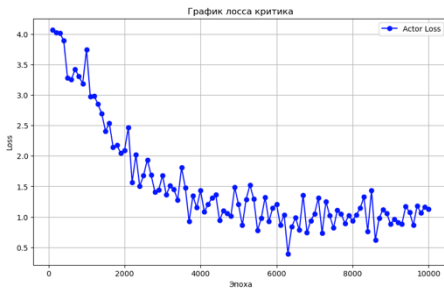


Архитектура нейронной сети



```
DockingAgent(
  (protein_encoder): ProteinEncoder3DCNN(
    (conv1): Conv3d(4, 16, kernel_size=(3, 3, 3), stride=(1, 1, 1), padding=(1, 1, 1))
    (bn1): BatchNorm3d(16, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv2): Conv3d(16, 32, kernel_size=(3, 3, 3), stride=(1, 1, 1), padding=(1, 1, 1))
    (bn2): BatchNorm3d(32, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv3): Conv3d(32, 64, kernel_size=(3, 3, 3), stride=(1, 1, 1), padding=(1, 1, 1))
    (bn3): BatchNorm3d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (pool): MaxPool3d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
    (fc): Linear(in_features=4096, out_features=128, bias=True)
  )
  (ligand_encoder): LigandGNN(
    (convs): ModuleList(
      (0): GCNConv(30, 64)
      (1): GCNConv(64, 64)
      (2): GCNConv(64, 128)
    )
  )
  (actor): Actor(
    (fc1): Linear(in_features=259, out_features=256, bias=True)
    (fc2): Linear(in_features=256, out_features=256, bias=True)
    (mean_head): Linear(in_features=256, out_features=3, bias=True)
    (log_std_head): Linear(in_features=256, out_features=3, bias=True)
  )
  (critic): Critic(
    (fc1): Linear(in_features=262, out_features=256, bias=True)
    (fc2): Linear(in_features=256, out_features=256, bias=True)
    (q_head): Linear(in_features=256, out_features=1, bias=True)
  )
)
```

Результаты



Результаты

Метрика	Значение
MAE (средняя абсолютная ошибка)	1.248
MSE (среднеквадратичная ошибка)	1.141
RMSE (корень из MSE)	1.117
MAPE (средняя абсолютная процентная ошибка)	22.3%
R^2 (коэффициент детерминации)	0.71

Спасибо за внимание!