

DRUG TARGET PREDICTION

Ashray Shetty



Problem with drug discovery?

- Drugs can bind to more than one target
- Sometimes the connection can be weak
- Only around 10% of candidate drugs are approved
- Half of drugs rejected due to lack of **effectiveness**



How to identify drug's target?

Based on structure

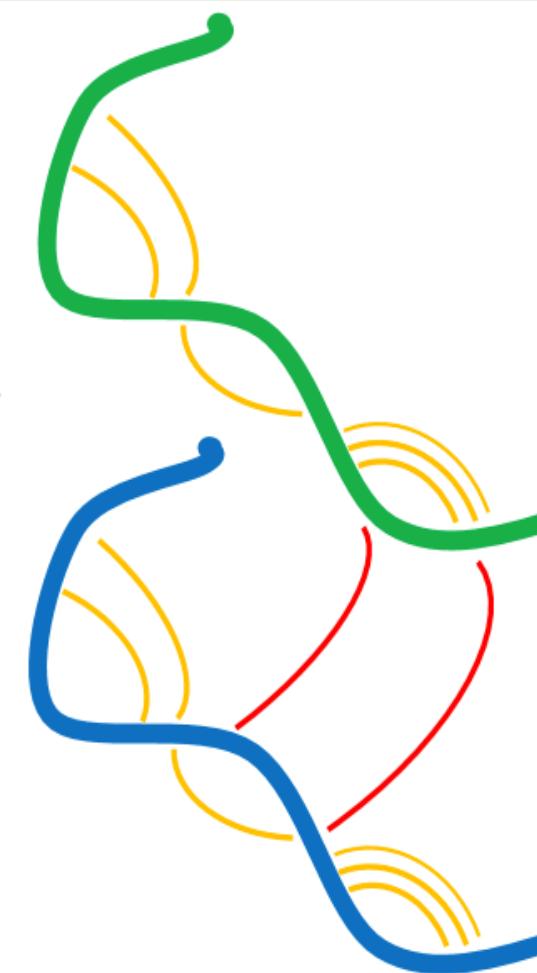
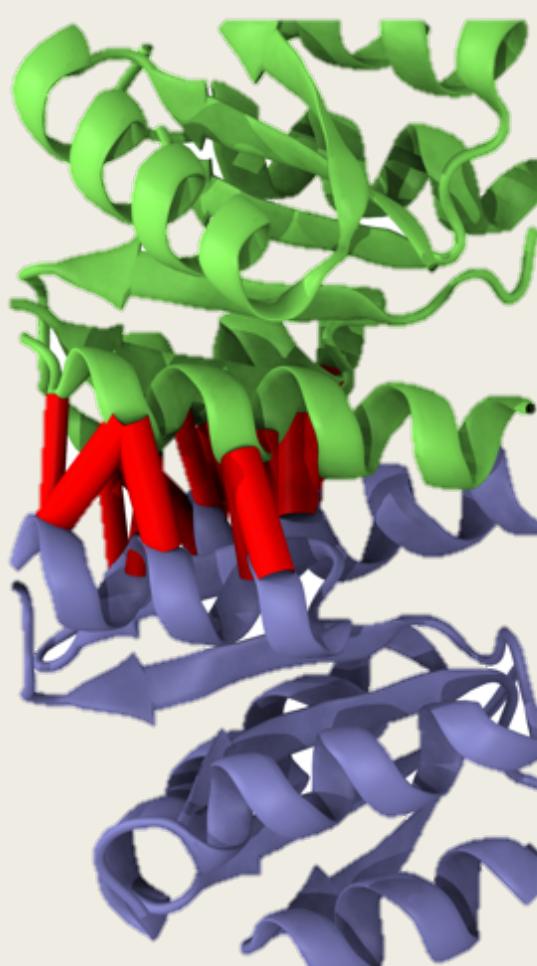
- very complex
- very expensive
- cannot be automated

Using chemistry

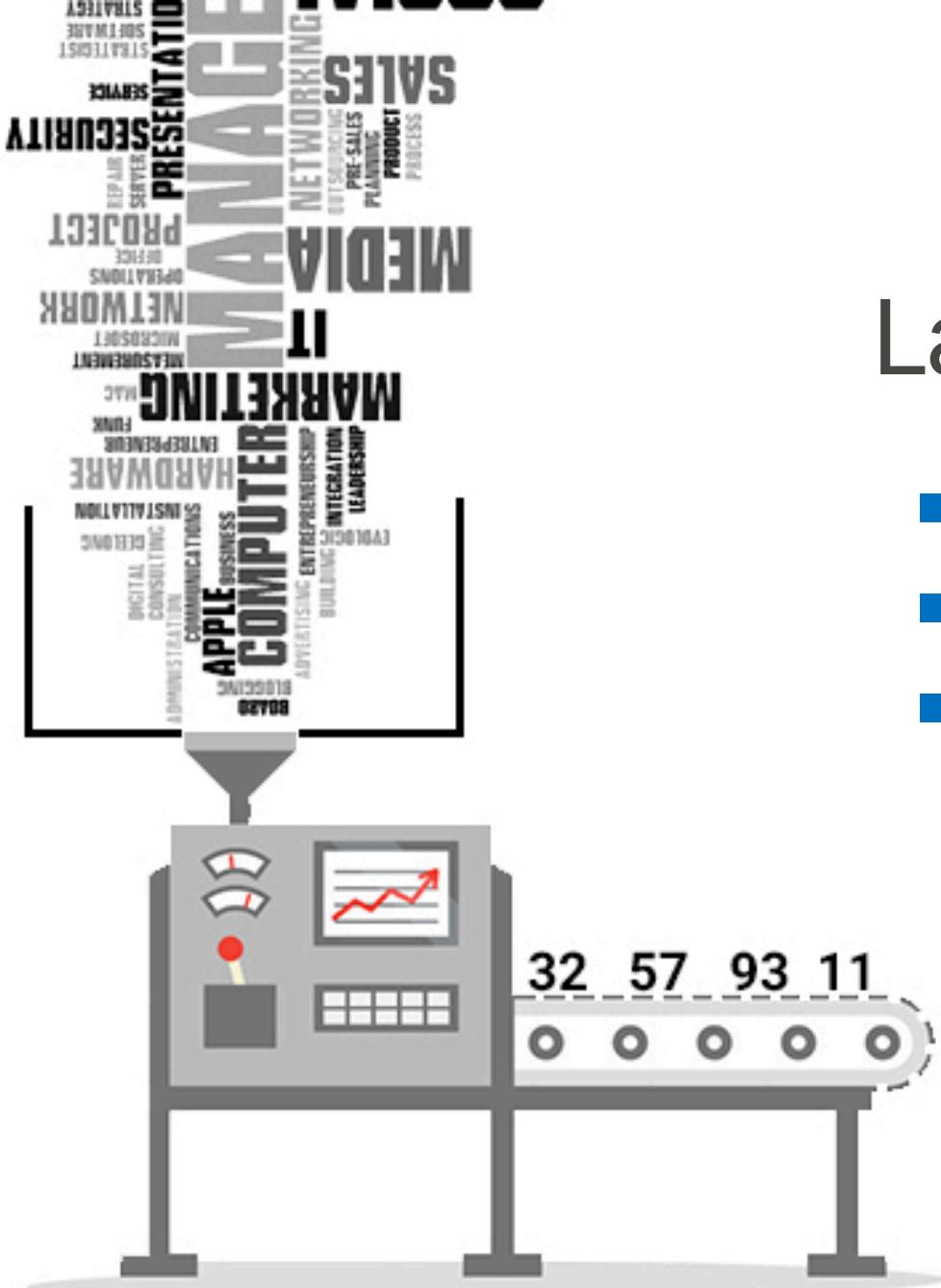
- Automatized
- but expensive
- and time-consuming



Can we use computational approach?



RIDHRLKNTDH
FLNGRLRDTDH
HERQETGELKH
KYRTRLTDLDH
RRAMEVGNLKH
TQKEELANLKH
KQQSEVENAKH
RLNQRADDLDH

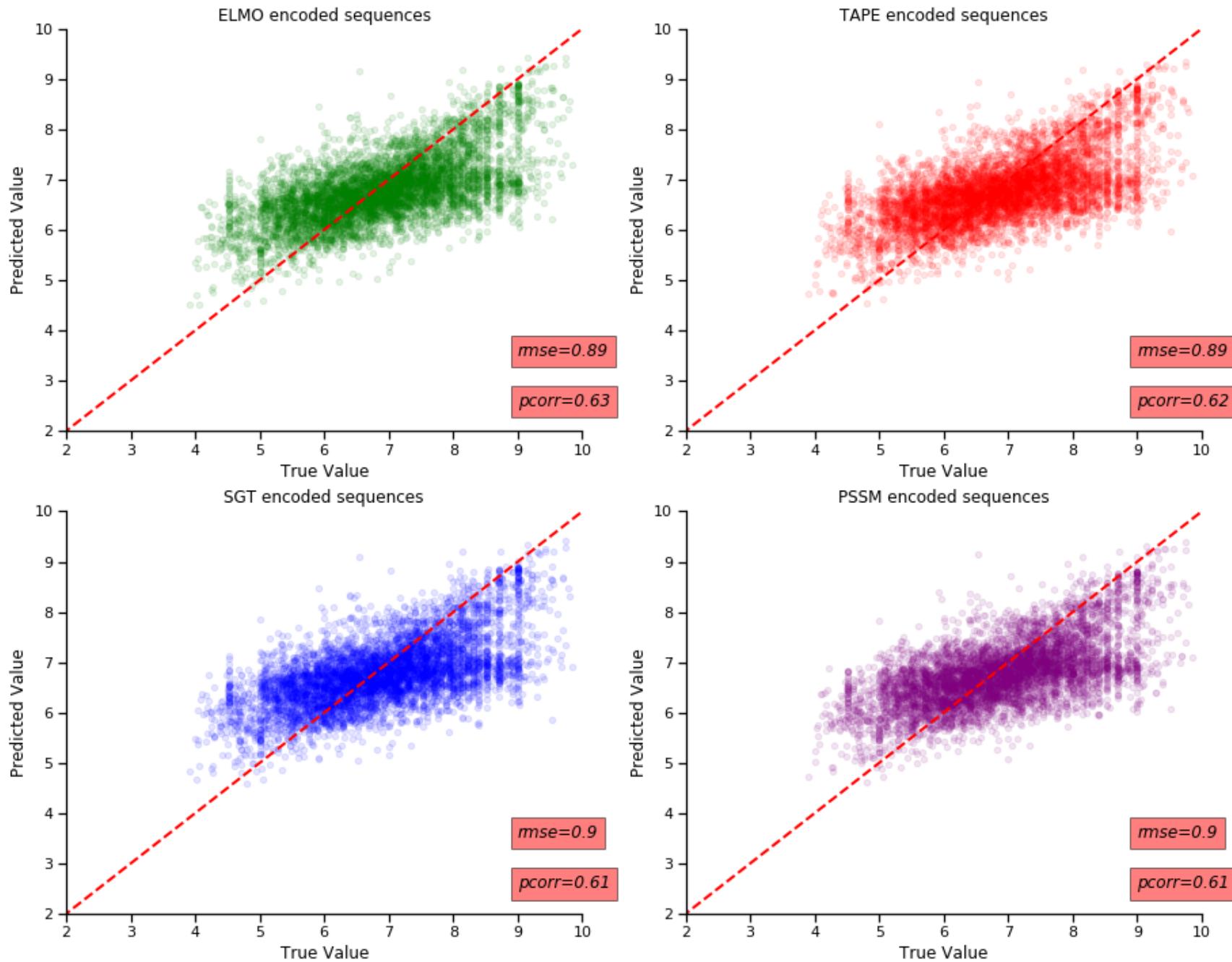


Language processing

- Turn letters to numbers
- Extract hidden meaning
- Useful for machine learning

Project overview:

- Dataset containing 89000 drug-target binding value
- Drugs have potential use as anticancer medication
- Used 4 different language processing methods for representing proteins



Which drug is best?

Drugs

'CHEMBL223361', 'CHEMBL182648',
'CHEMBL1773509', 'CHEMBL3219238',
'CHEMBL432116', 'CHEMBL3909781',
'CHEMBL255955', 'CHEMBL489684',
'CHEMBL3601715', 'CHEMBL3402367',
'CHEMBL230809', 'CHEMBL236924',
'CHEMBL310724', 'CHEMBL3809912',
'CHEMBL3661430', 'CHEMBL179753',
'CHEMBL378347', 'CHEMBL3939230',
'CHEMBL3360333', 'CHEMBL2158850'

Proteins

'P35968', 'Q16539', 'Q05397', 'P43405',
'O60674', 'P10721', 'O14965', 'P49841',
'P52333', 'P12931', 'Q96GD4', 'P36888',
'P06239', 'P11802'

compound_id	P35968	Q16539	Q05397	P43405	O60674	P10721	O14965	P49841	P52333	P12931	Q96GD4	P36888	P06239	P11802
CHEMBL223361	6.9	6.8	5.6	6.8	6.4	7	6.4	6.1	6.4	5.7	7	7.3	6.1	6.1
CHEMBL182648	7.1	6.9	5.5	7	6.8	7.1	6.4	6.6	6.6	5.6	6.8	7.2	6.1	6.6
CHEMBL1773509	6.3	7	5.6	6.7	7	6.6	6.2	6.4	6.4	6	6.3	6.8	6.4	6.2
CHEMBL3219238	6.7	7.1	5.7	7	6.5	6.9	6.4	6.8	6.5	6	6.8	7.1	6.7	6.2
CHEMBL432116	6.6	6.8	5.7	6.6	6.6	6.7	6.8	6.8	6.4	6.1	6.6	6.6	6.5	5.6
CHEMBL3909781	6.8	6.7	5.8	7.5	7.1	7	6.8	6.7	6.6	6.7	6.7	7	6.7	6.8
CHEMBL255955	7	6.7	5.7	6.6	6.4	7	6.3	6.2	6.5	5.8	6.9	7.1	6.3	6.1
CHEMBL489684	6.8	6.9	6.5	7.4	7.1	7.5	7	6.5	6.6	6.3	6.9	7.6	6.7	7.3
CHEMBL3601715	6.8	6.9	6.8	7	6.8	6.9	7	6.8	7.1	6.7	7.1	6.9	6.7	6.9
CHEMBL3402367	6.7	6.2	6	5.8	6.7	6.8	6.9	6.4	6.4	6.3	6.7	6.8	6.5	6.5
CHEMBL230809	5.8	6.5	5.4	6.3	6.7	6.1	5.8	6.1	6.1	5.2	6.3	6.1	5.2	5.9
CHEMBL236924	6	6.6	5.4	6.3	6.4	6.2	5.9	6.7	6.4	5.7	6.1	6.2	6	5.9
CHEMBL310724	6.1	6.5	5.5	6	6.1	6.2	6.1	6.4	6.1	6	6.2	6.3	5.8	6.1
CHEMBL3809912	6.6	6.5	5.5	6.4	6.3	7	6.2	6.3	6.4	5.5	6.8	7.1	6.1	6.2
CHEMBL3661430	7	7.6	8.7	7.1	7	7	7	6.8	6.9	7.3	7.2	7.4	7.3	7.1
CHEMBL179753	7.1	6.8	5.6	6.9	6.6	7.2	6.5	6.4	6	6.3	6.9	7.3	6.5	6.3
CHEMBL378347	6.4	6.4	5.3	6	6.5	6.1	5.9	6.6	6.4	5.3	6	6.2	5.8	5.8
CHEMBL3939230	6.9	7.2	6.3	7.3	7.2	7.2	6.8	6.6	7.9	6.4	6.8	7.1	6.6	7
CHEMBL3360333	6.2	6.4	5.4	6.9	6.7	6.7	6.2	6.3	6.4	6	6.3	6.6	6.6	6.6
CHEMBL2158850	6.3	6.6	5.9	6.7	7	6.5	6.4	6.7	6.3	6.5	6.5	6.7	6.6	6.4

Benefits of using AI based screening

Increase likelihood of success to achieve a pre-clinical candidate and reduce the time taken due to lower number of Med Chem design cycles and bioassays

Can classify drugs and plan their activity measurements accordingly

Allows us to fill the blank in our understanding of drug target activity space

Better able to detect and avoid side effects associated with drug

Thank you