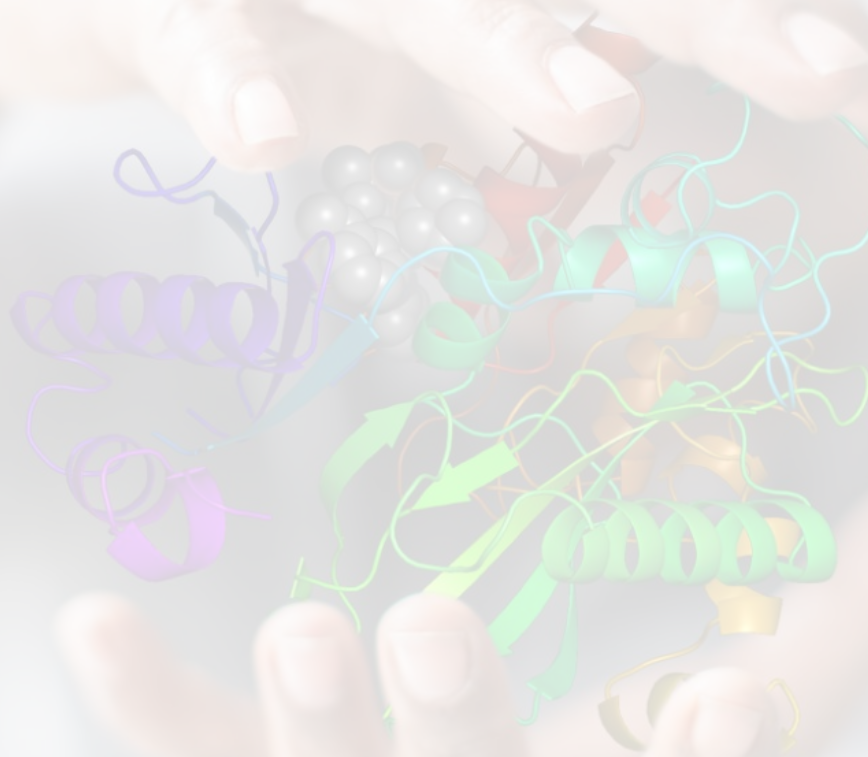


Identifying drug-like molecules that bind to and act through GPCRs



Discovering potential therapeutic targets

Diseased animal



Knockout protein

Healthy animal



Discovering potential therapeutic targets

Diseased animal



Knockout protein

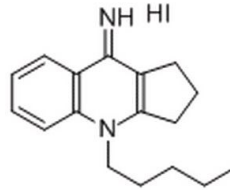
Healthy animal



There are currently no viable methods to safely knockout proteins in humans

Discovering potential therapeutic targets

Diseased animal



Protein inhibition



Healthy animal



Instead inhibit the protein to have a similar effect.

Reduce compounds in high-throughput screen

Molecules in a high-throughput screen:

200,000 to 1,000,000+

Reduce compounds in high-throughput screen

Molecules in a high-throughput screen:

200,000 to 1,000,000+

Cost to screen:

\$0.45/molecule

Reduce compounds in high-throughput screen

Molecules in a high-throughput screen:

200,000 to 1,000,000+

Cost to screen:

\$0.45/molecule

Cost for one screen:

\$90,000 to 450,000+

Reduce compounds in high-throughput screen

Molecules in a high-throughput screen:

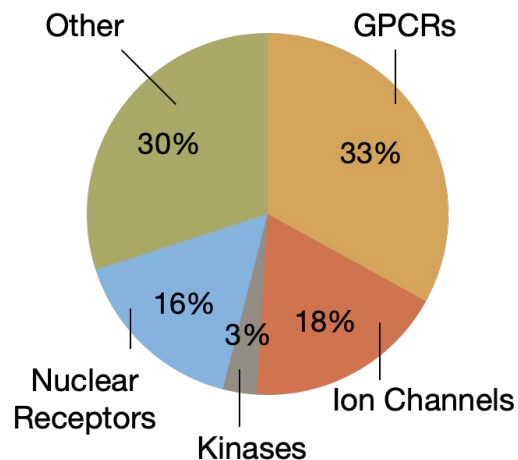
200,000 to 1,000,000+

Cost to screen:

\$0.45/molecule

Cost for one screen:

\$90,000 to 450,000+



~33% of FDA-approved drugs target GPCRs

Reduce compounds in high-throughput screen

Molecules in a high-throughput screen:

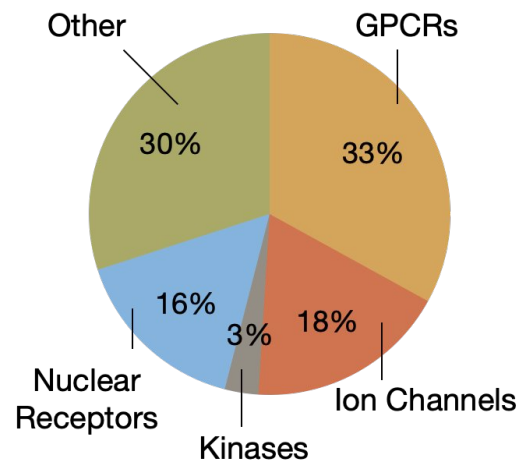
200,000 to 1,000,000+

Cost to screen:

\$0.45/molecule

Cost for one screen:

\$90,000 to 450,000+



~33% of FDA-approved drugs target GPCRs

Classify molecules to screen as GPCR interactor or not

Approach to filter molecules for HTS

Data

Molecules and
their targets

Molecular
properties

Target protein
family

Approach to filter molecules for HTS

Data

Molecules and
their targets

Molecular
properties

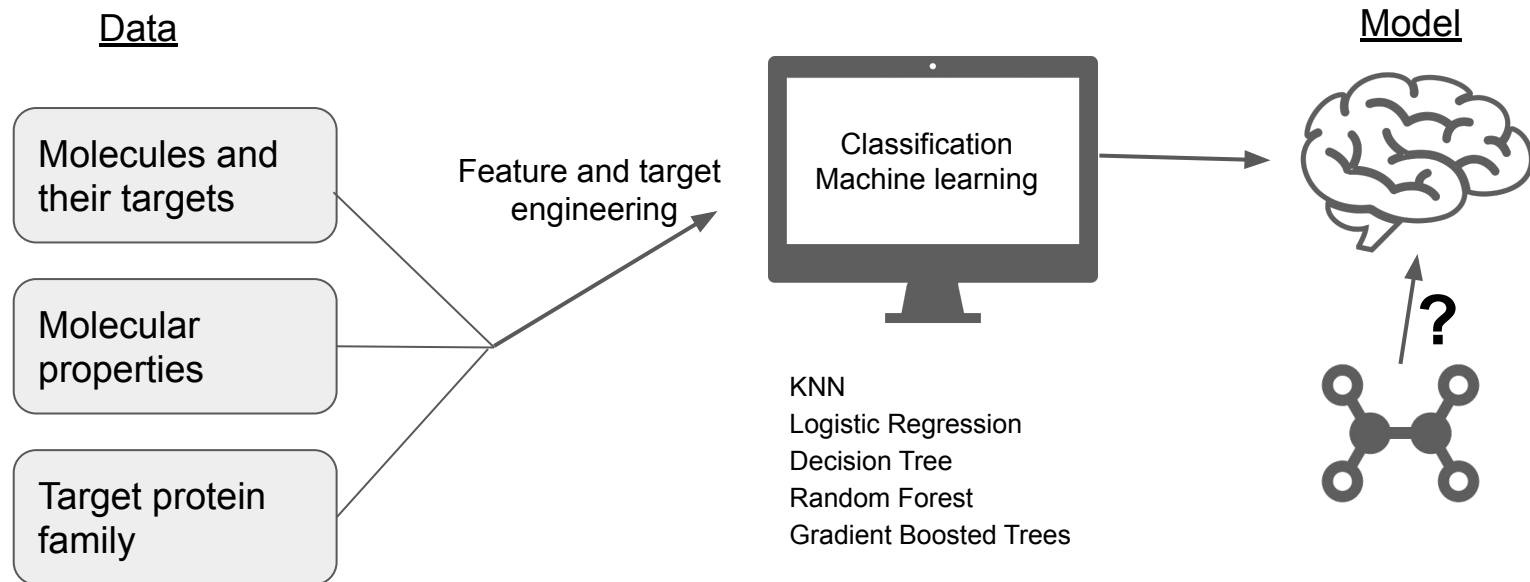
Target protein
family

Feature and target
engineering

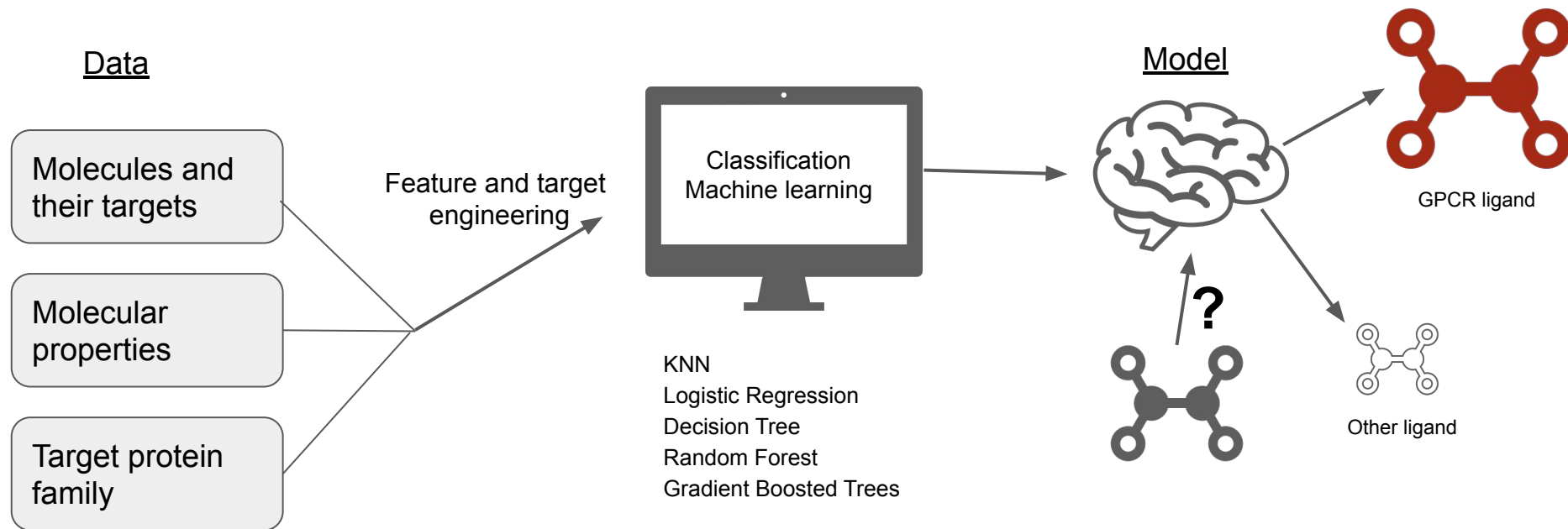
Classification
Machine learning

KNN
Logistic Regression
Decision Tree
Random Forest
Gradient Boosted Trees

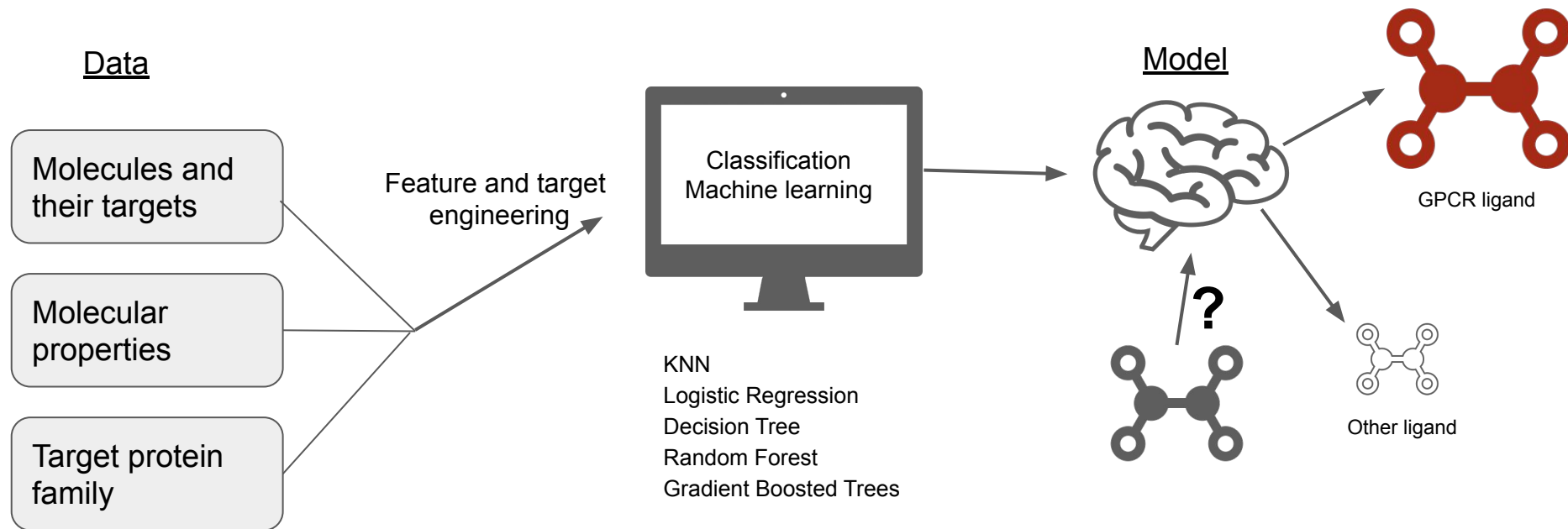
Approach to filter molecules for HTS



Approach to filter molecules for HTS

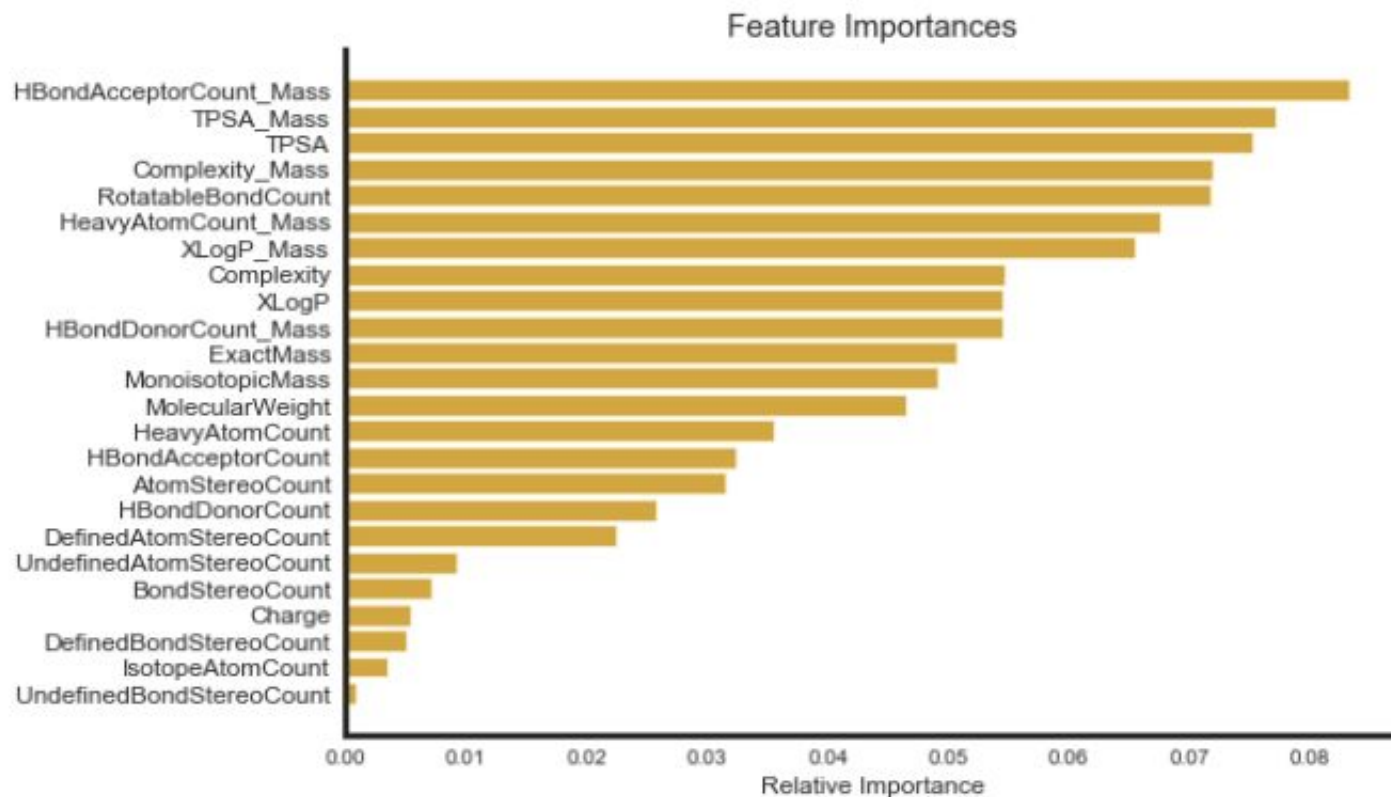


Approach to filter molecules for HTS

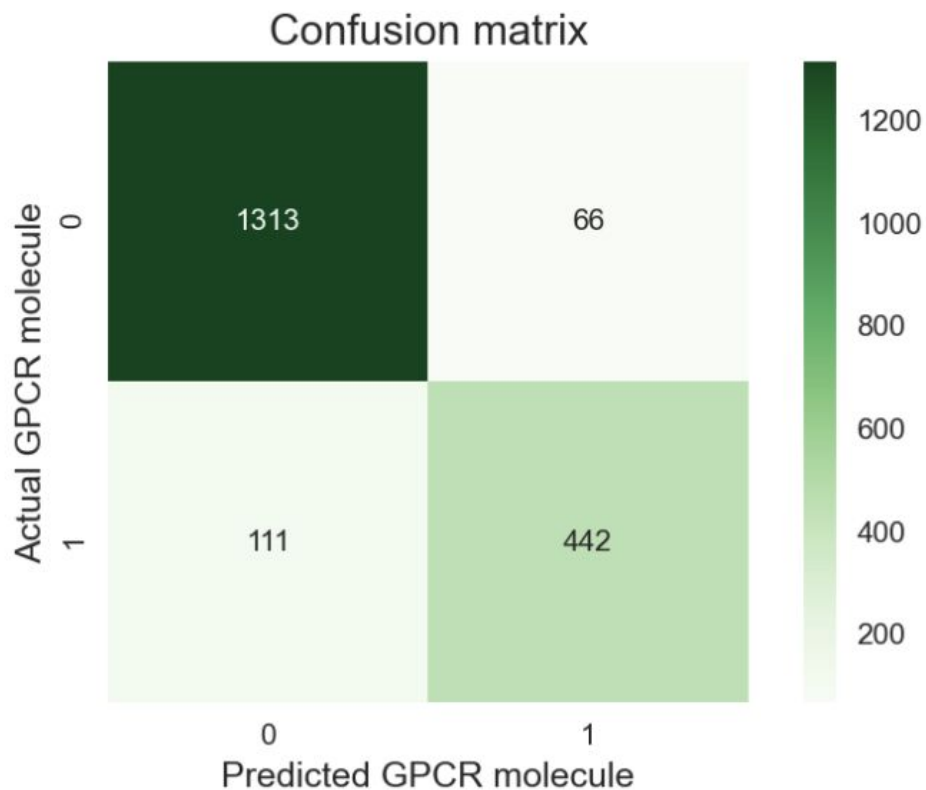


Goal is to have 95% recall and correctly identify 50% of other ligands

Feature Importance



Model performance



Conclusions

- High throughput screen with
- 500,000 molecules
 - same GPCR ligands:
non-GPCR ligands as dataset

Recall threshold	Molecules filtered	Money saved
99%	216,000	\$100,000
95%	300,000	\$135,000
90%	352,000	\$156,000

Use machine learning to design better high throughput screens