

PREDICTING ONTOLOGY OF UNKNOWN NATURAL PRODUCTS

Meir Alelov

SUMMARY

- Natural product ontology classification
- Predictions provide key insights
- Drug discovery pipeline efficiency

PRESENTATION OUTLINE



Business Problem & Data

Modeling & Evaluation



Conclusions & Next Steps

BUSINESS UNDERSTANDING

Problem

- \$2.6 billion, 10 years
- Plethora of methods & parameters

Solution

- Natural Products
- Ontology classification
- Ontology knowledge = pipeline efficiency

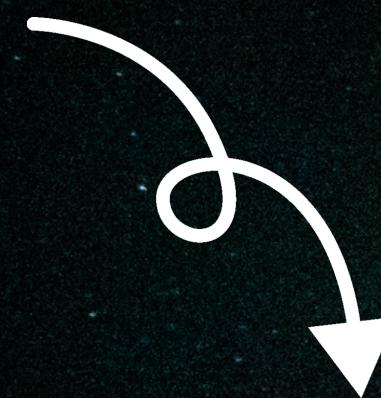
DATA UNDERSTANDING

- PubChem and ChEBI
- Natural products
- One Role Type



MORE INFO

10,488 compounds, 18 features

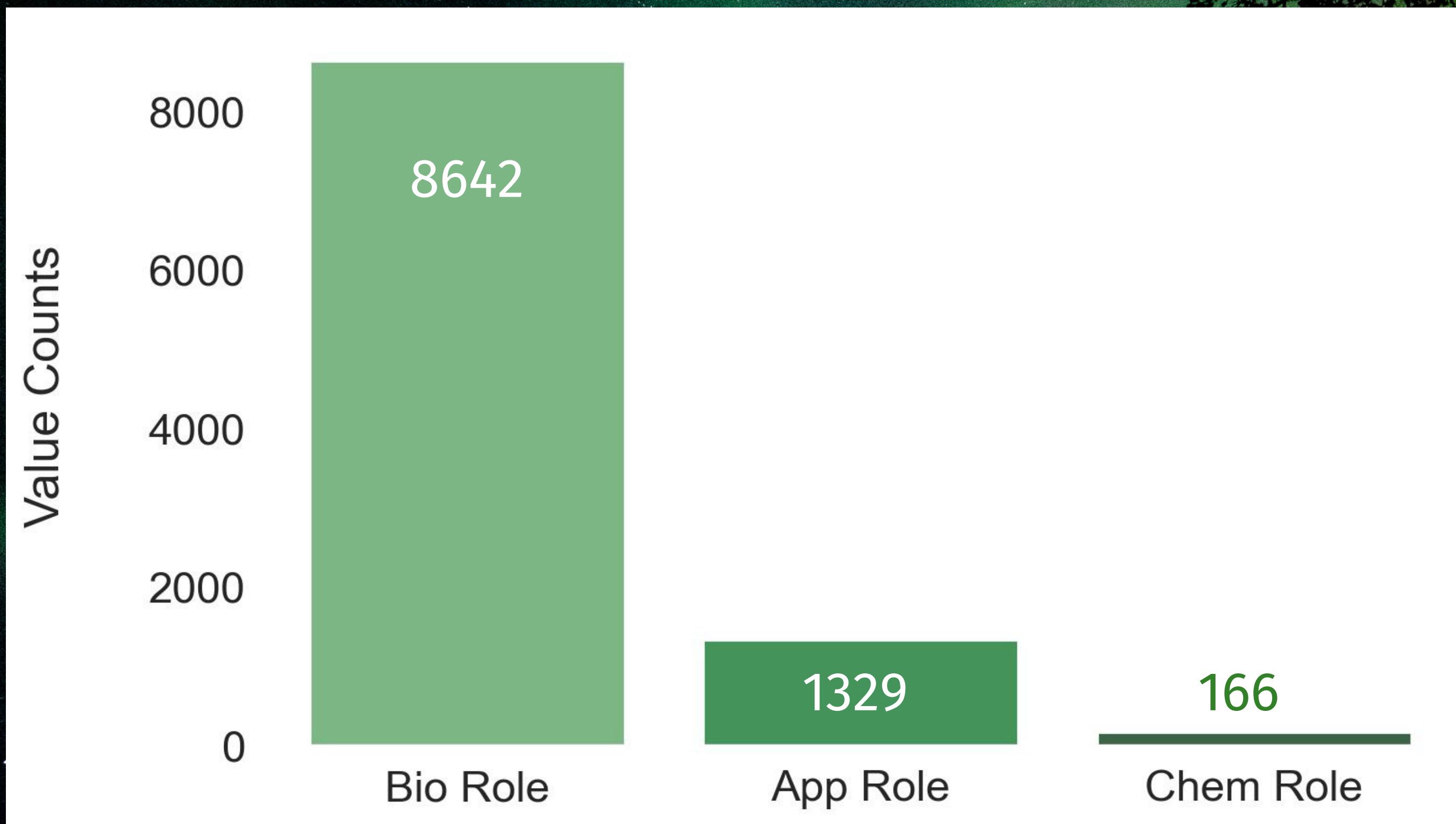


Data: Chemical & physical properties

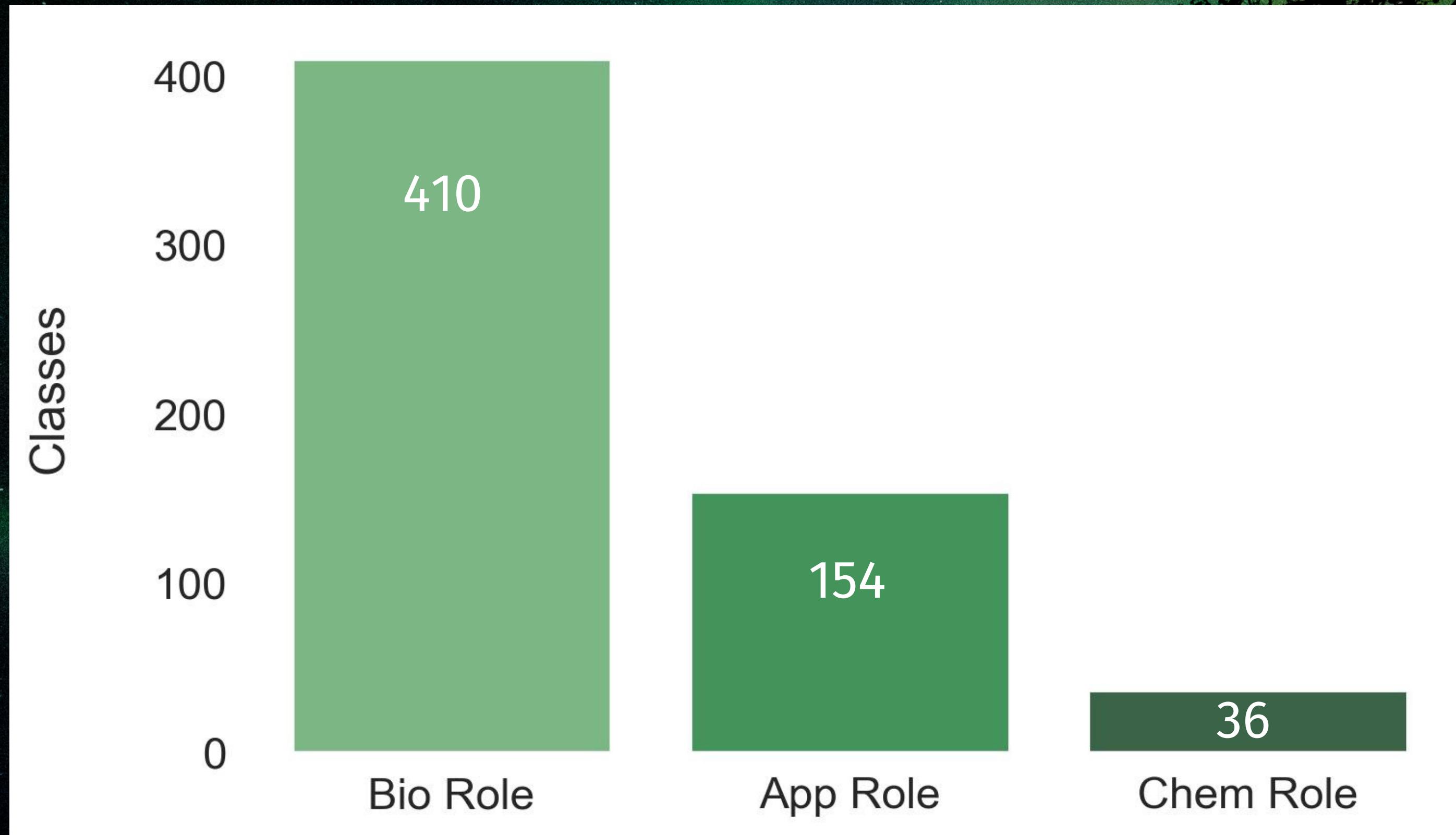


Targets: Role Type & Role Name

ROLE TYPE TARGETS

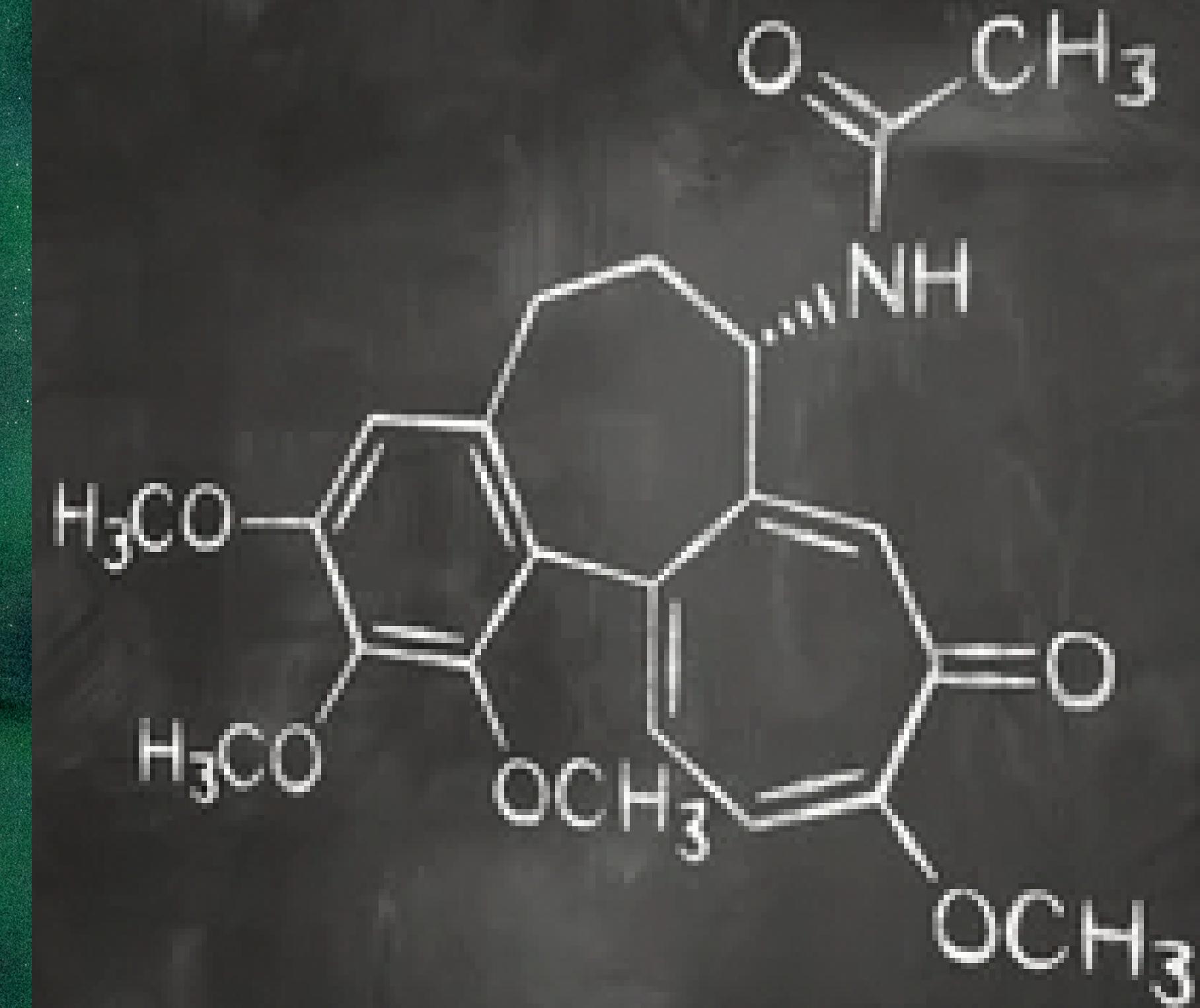


ROLE NAME TARGETS

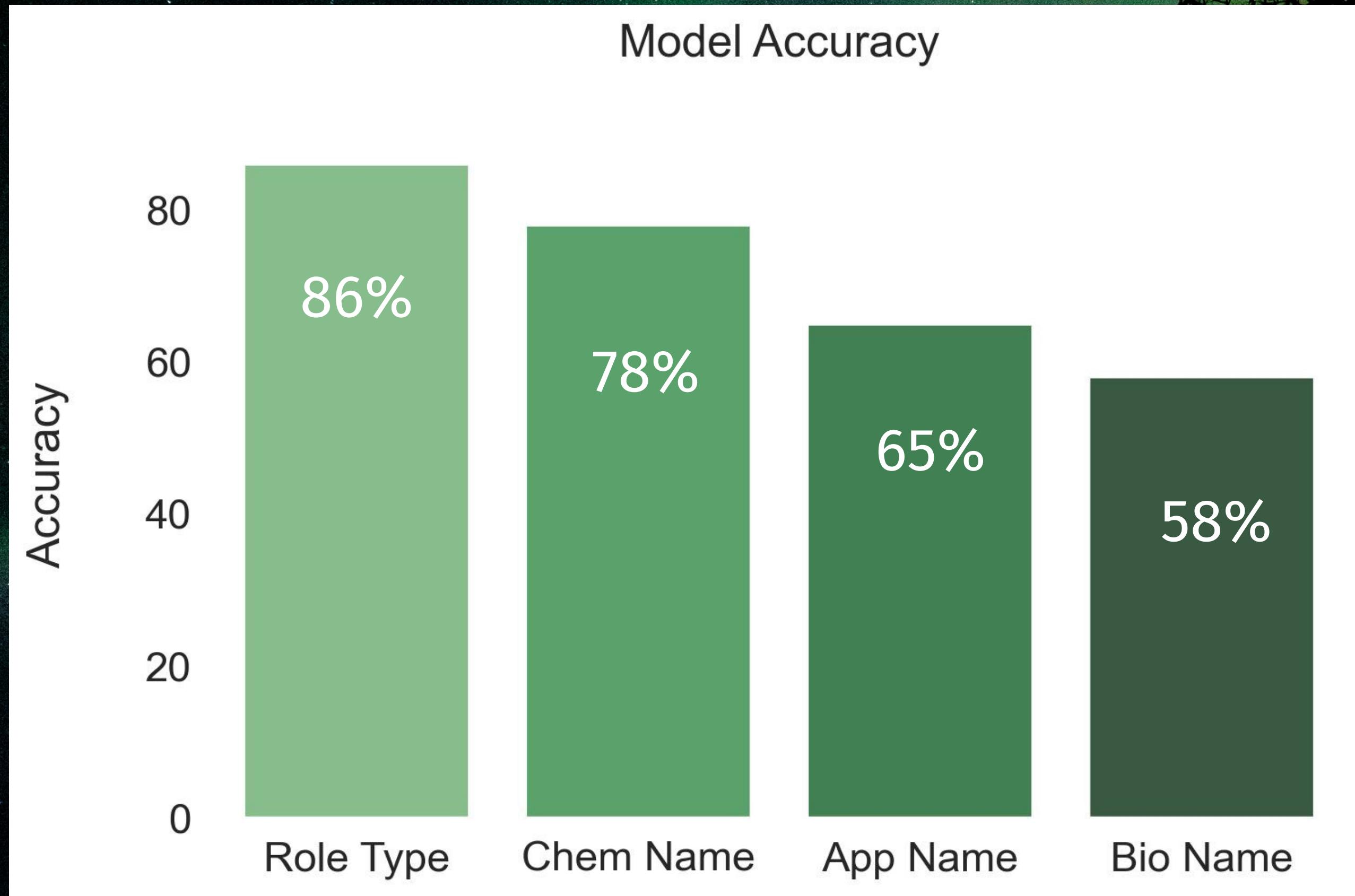


METHODS & MODELING

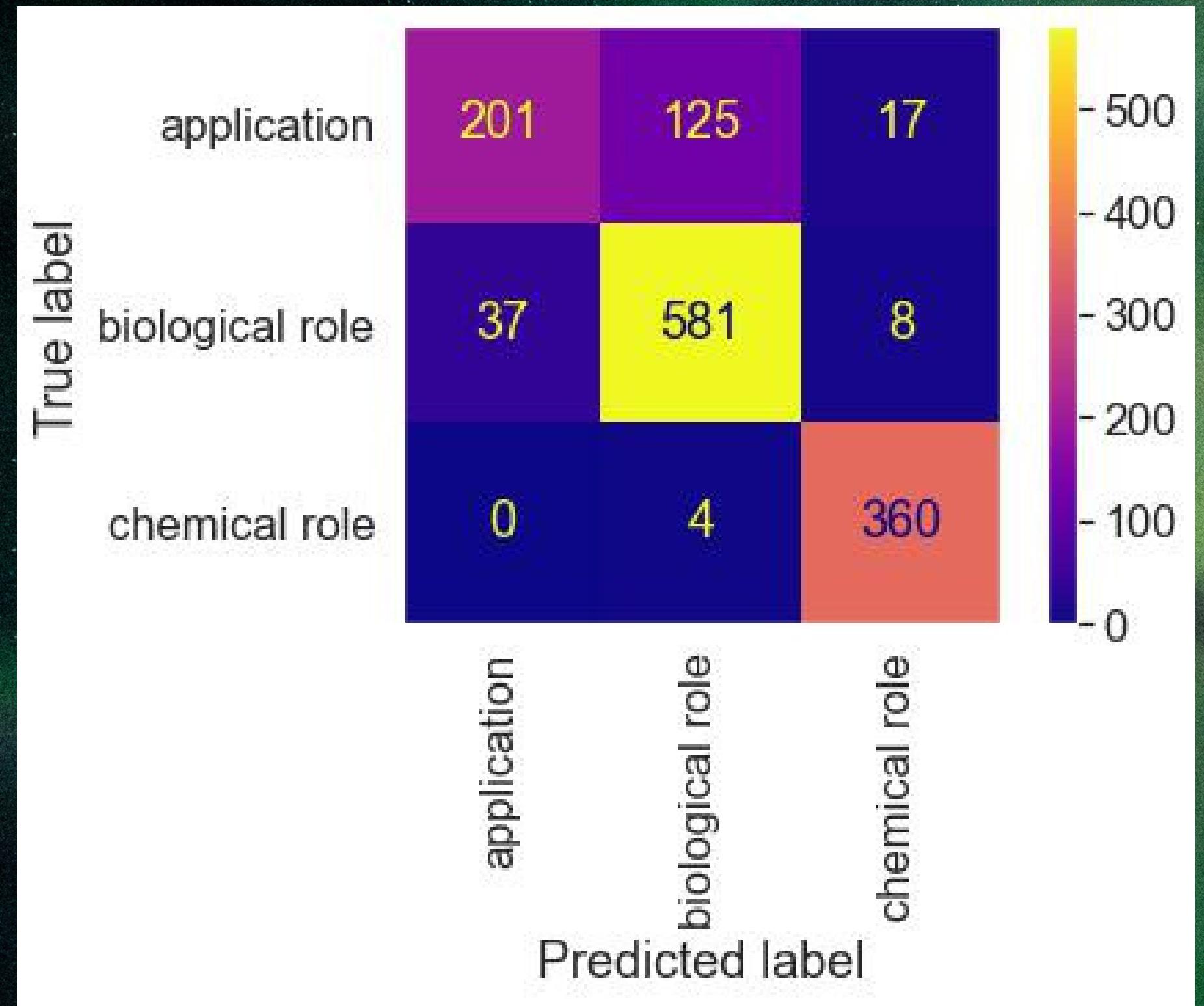
- ChemParse
- Class Imbalance
- Random Forest Classifier



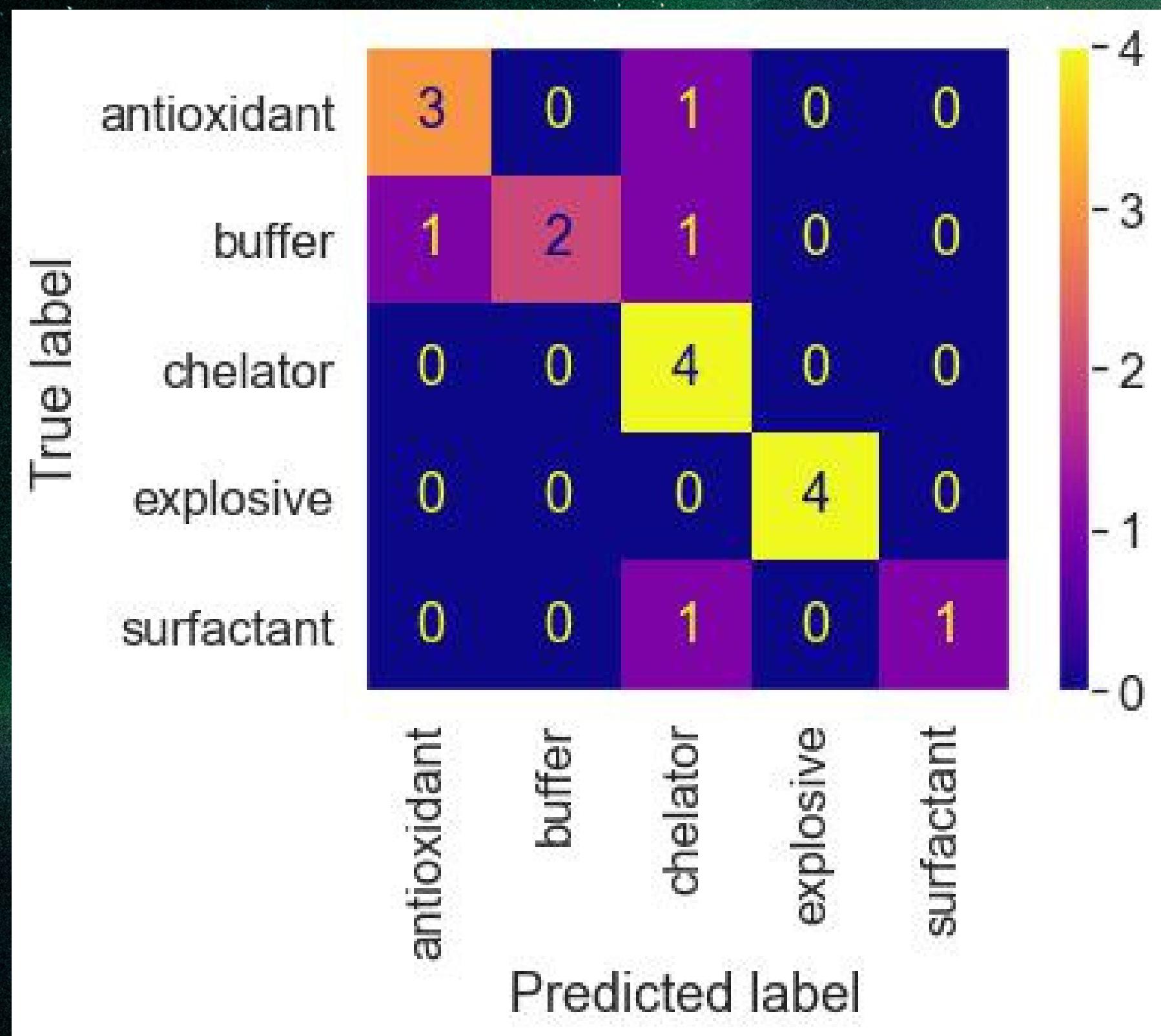
RESULTS & EVALUATION



ROLE TYPE



CHEM NAME



Chem Name: 78%

APP DEMO

Natural Product Classifier

This app takes in a newly discovered naturally occurring compound

It then predicts the type of role it has in a drug pipeline, as well as the name of that role

Molecular Formula

C₁₈H₁₀O₅

Molecular Weight

306.30

XLogP

3.00

Exact Mass

306.05

Monoisotopic Mass

306.05



Home Advanced Search Browse Documentation Download Tools About

Chemical Entities of Biological Interest (ChEBI) is a freely available dictionary of molecular entities focused on 'small' chemical compounds.

Search i

Search for ★★★ only All in ChEBI

Example: [iron*](#), [InChI=1S/CH4O/c1-2/h2H,1H3](#), [caffeine](#)

[Advanced Search](#) | [About ChEBI](#)

Downloads

News

SDF files ChEBI
provides its chemical

Tweets by @chebit i

This website requires cookies, and the limited processing of your personal data in order to function. By using the site you are agreeing to this as outlined in our [Privacy Notice](#) and [Terms of Use](#).

FEATURE IMPORTANCE

- Complexity
- Carbon
- Hydrogen
- TPSA
- Rotatable Bond Count

RECOMMENDATIONS

- Mode of Action
- Avoid Unnecessary Testing
- Testing Parameters



NEXT STEPS

- Class Combinations
- More Data
- Model Complexity

Thank you!

Meir Alelov
[GitHub](#)
[Email](#)



Resources

<https://www.u-bordeaux.com/Education/Study-offer/Masters-in-English/Analytical-Chemistry-for-Drugs-and-Natural-Products>

<https://rhochi.org/natural-drug-research-natural-organic-and-scientific-extraction-in-glassware-alternative-green-herb-medicine-natural-skin-care-beauty-products-laboratory-and-development-concept/>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4211845/>

<https://www.technologynetworks.com/drug-discovery/articles/target-identification-validation-in-drug-discovery-312290>

<https://www.nature.com/articles/s41573-020-00114-z#Sec6>

<https://pgsolx.com/Press/?p=2739>