EXECUTIVE SUMMARY

118th Congress Project

This project uses logistic regression to predict whether a bill introduced in the 118th Congress will become a law.

As of 1 June 2024, 15,366 bills have been introduced in the 118th Congress but only 64 have become law! This means there is a severe class imbalance in the data. To mitigate this, the algorithm was trained and used on four subsets of the data:

- All bills (15,366).
- Bills that passed in the House (539). Given that a bill passed in the House, will it become a law?
- Bills that passed in the Senate (189). Given that a bill passed in the Senate, will it become a law?
- Bills that passed in both chambers (81). Given that a bill passed in both chambers, will it become a law?

Results

Train.

Trained on	Baseline Accuracy	Accuracy	Precision	Recall
All bills	99.6%	99.8%	100%	57.1%
Passed in House	90.7%	97.7%	80%	100%
Passed in Senate	66.7%	85.7%	83.3%	71.4%
Passed in both	33.3%	83.3%	80%	100%

Test. We chose to test with the algorithm that was trained on the bills that passed in the House. Here is how the model performed:

Baseline Accuracy	Accuracy	Precision	Recall
94.3%	97.7%	80%	57 1%

Implications

When training the KPIs looked pretty good, but when tested the model was likely overfit due to the severe class imbalance and due to the large number of features. The next step is to work some more on dimension reduction and pinpoint which features lead to a bill becoming law.