Forecasting model could predict which bills get passed

August 7 2018, by Steve Tally

Purdue University's professor of political science Eric Waltenburg and assistant professor of computer science Dan Goldwasser have teamed up to design a forecasting model that uses data science to determine the votes of state legislators. Credit: Purdue University

When the United States pulled out of the Paris Climate Accord last year,

20 states—including New York and California—resolved to form a new, nonfederal agreement to address climate change and greenhouse gas emissions.

According to Purdue University professor of political science Eric Waltenburg, many meaningful policy decisions such as this take place at the state level but are largely ignored by the public. He is collaborating with Dan Goldwasser, assistant professor of computer science, to help people better understand the impact of decisions made at the state level. Together they lead a project focused on developing a forecasting model to predict state legislator voting behavior.

"If our project works like I'm hoping it does, it will open up the policy-making process and people will have a better sense as to what legislative outcomes may be," Waltenburg said. "It would demystify the state legislative process for the public."

State legislatures are assemblies of elected members who vote on laws and policy for their state. In a roll call vote, each legislator's name is called, and they must indicate on record whether they are for or against a particular piece of legislation. Using public sources such as Twitter, political blogs, newspaper accounts and historical roll-call data from the past decade, computer scientists working with Waltenburg will use machine learning techniques to predict which way a legislator may vote on a particular bill.

"What we're trying to do is identify key forces – I refer to them as revealed preferences – among legislators and use those preferences to predict outcomes," Waltenburg said. "We have huge amounts of historical roll-call data at the state level, and with that data we're able to produce relational matrices for how each legislator votes in relation to every other member of the legislature."

Ultimately, they aim to produce a website where people could go to predict the outcome of a particular piece of legislation. Goldwasser added that the impact of decisions made at the state level are more significant than people think, despite the fact that they will not get the same coverage as those made at the national level.

"We're really hoping that this would be a public service that allows people to hold their representatives accountable and at least make more informed decisions when they go to the polls," Goldwasser said.

By the end of the summer, their project will have expanded beyond the state of Indiana to Oregon and Wisconsin, where they will test their forecasting model on the lower-level chambers of those states. Ultimately, they would like to transform their model from static to dynamic, bringing in real-time data from Twitter and other online sources so that the model can adapt the relational matrix as the revealed preferences come in.

"One of the questions we are asking is, how can we use online resources such as Twitter or news coverage to dynamically update the model that we have about those preferences?" Goldwasser said. "When how a state legislator is going to behave is not explicitly mentioned, we can learn to infer their behavior."

The team has a static model for the Indiana House but would ultimately like to develop a forecasting model for all states and their chambers. Their project was one of eight selected by the Purdue's Integrative Data Science Initiative to be funded for a two-year period. The initiative will encourage interdisciplinary collaboration and build on Purdue's strengths to position the university as a leader in data science research and focus on one of four areas: health care; defense; ethics, society and policy; and fundamentals, methods, and algorithms. The research thrusts of the Integrative Data Science Initiative is hosted by Purdue's Discovery Park.

"Although the idea of predicting legislative votes is a pretty conventional political science topic, Dan's expertise allows us to develop a forecasting model that goes well beyond where most political scientists have gotten," said Waltenburg. "The project we are working on is a great example of how Purdue's resources and faculty help each other implement ideas and collaborate across campus and disciplines."

Provided by Purdue University